مبادئ علم الأحصاء

مبادئ علم الإحصاء

الدكتور
أحمد فوزى ملوخية
وكيل المعهد العالى للسياحة والفنادق
الإسكندرية

2005

مكتبة بستان المعرفة
طباعة ونشر وتوزيع الكتب
☎: ٠٤٥/٢٢٢٤٢٢٨
٠١٢١١٥١٢٣٧&٠١٢٣٥٣٤٨١٤

بسم الله الرحمن الرحيم

تقديم

لاشك أن مصر تواجه تحديات كثيرة وخاصة فى ظل الظروف البيئية الداخلية والخارجية ولاشك أيضا أن التركيز على العلم والمعرفة والتكنولوجيا أصبح واجبا وحتميا لمواجهة هذه التحديات.

لذلك كان هذا الكتاب محاولة بسيطة الى أبعد الحدود مع الاحتفاظ بالموضوعية فى نفس الوقت. وقت قدم موضوعات فى طبيعة علم الإحصاء والبيانات والمعلومات وعرضها بيانيا وجدوليا وذلك إ نطلاقا من أن عملية تحليل البيانات هى عملية تحرى Dective work عن طريق العدو الأعداد والأشكال والتى تقع مسئوليتها الأولى والاخيرة على عاتق الباحث ويكون دور الحاسبات الالكترونية هو معاونة الباحث على تنفيذ استراتيجيات التحليل التى توصل إليها بدرجة أكثر فاعلية ومرونة.

ولا ندعى أن الكتاب يخلو من نقائص فليس فى وسع أى باحث وما كانت مقدرته العلمية أن يصل بدراسته إلى درجة الكمال، فهو لله وحده، ولكنها محاولة لاتزال بحاجة إلى مزيد من التدعيم والمعالجة المستفيضة التى أرجو أن تتاح لنا فرصة تحقيقها فى المستقبل القريب.

وندعو الله أن يوفقنا لما فيه قصد السبيل وأن يحقق هذا الكتاب الغرض من إصداره.

د/ احمد فوزي ملوحية

الباب الأول

المفاهيم الإحصائية

المفاهيم الإحصائية

فى الحقيقة أن كلمة الاحصاء باللغة الانجليزية Statistics تستخدم فى حالة المفرد وحالة الجمع . ففى حالة الجمع فان :

الاحصاء (ونطلق عليها فى العربية الاحصاءات) يمثل البيانات الخام ذاتها والحقائق الرقمية التى تمثل نتائج القياس او التجريب او النشاط .

وفى حالة المفرد فان الاحصاء يتعامل مع جمع وتحليل وتفسير البيانات واتخاذ القرار . ومن ثم فان :

علم الاحصاء يمثل ذلك الفرع من الرياضيات الذى يسهل ويساعد فى اتخاذ القرارات الحكيمة او الرشيدة فى مواجهة عدم التاكد ، ومن ثم يستخدم الأدوات والوسائل الفنية لجمع البيانات الدقيقة عن المجتمع تحت الدراسة وعرضها بصورة فعالة وتحليل وتفسير دقيق للمعلومات الرقمية حتى يمكن اتخاذ القرارات الحكيمة .

ان المديرين والإدارات الحديثة الواعية والاقتصاديون مثلا يعتمدون دائما على المدخلات من التحليل الاحصائى للمساعدة فى اتخاذ القرارات . ان مشكلة اتخاذ القرار الإدارى أو مشكلة اتخاذ القرار الاقتصادى تتعلق دائما بظاهرة أو ظواهر حقيقية فى المجتمع تحت الدراسة

ان الاحصاء فى سنواته الأولى كان يستخدم بمعرفة الاحصائيين بهدف جمع البيانات الاساسية التى تعتمد عليها الدولة للتعرف على خصائص السكان مثل البيانات عن المواليد والوفيات وعدد السكان لاغراض التجنيد فى الجيش وكذلك اهتمام الدولة بجمع بيانات عن الصادرات والواردات والدخول والانفاق بغرض معرفة حجم التجارة الخارجية ولفرض الضرائب ثم عرض هذه البيانات فى صورة مختصرة كالجداول والاشكال البيانية واستخدام بعض المقاييس الرقمية البسيطة بدلا من النظر الى كل البيانات الخام فى حجمها الاصلى لتفهم معنى هذه البيانات وهو ما كان يعرف بالاحصاء الوصفى .

ولذلك فان :

الاحصاء الوصفى عبارة عن ذلك الفرع من علم الاحصاء الذى يهتم بتجهيز البيانات التى جمعت بهدف تلخيص او وصف خصائص او طبيعة البيانات بدون محاولة استنتاج اى شئ بخلاف البيانات نفسها .

وبالرغم من أهمية الاحصاء الوصفى وزيادة رقعة استخدامه ، الا أن ضخامة البيانات عن المجتمعات تحت الدراسة ومحاولة جمعها يصبح أمرا عسيرا اذا لم يكن مستحيلا لأسباب كثيرة تتعلق بالتكاليف والوقت والدقة الى غير ذلك . وبهذه المناسبة فان لفظ " مجتمع " عند الاشارة اليه يتبادر الى كثير من الناس غير المتخصصين انه يعنى مجموعة السكان التى تعيش فى بلد معين . الا ان ذلك امرا خاطئا . ولذلك وجب ايضاح معنى لفظ مجتمع " population ولذلك فان :

المجتمع هو مجموعة كل المشاهدات المحتملة عن خاصية معينة لظاهرة او مجموعة من الظواهر التى لها خصائص مشتركة .

فمثلا تمثل دخول جميع العاملين فى شركة معينة مجتمع دخول العاملين أو ان تفضيلات جميع المستهلكين لسلعة معينة يمثل مجتمع تفضيلات المستهلكين أو ان خصائص جميع الطيور فى سماء مصر تمثل مجتمع خصائص الطيور أو اوزان الحيوانات أو الحالةالزوجية لمجتمع العاملين فى مشروع معين فان كل هذه الامثلة تمثل مجتمعات الدراسة .

ولصعوبة الحصول على بيانات المجتمعات تحت الدراسة للأسباب السابق ايضاحها فاننا نلجأ الى اختيار او سحب عينات Samples من هذه المجتمعات لمعرفة خصائص هذه المجتمعات . ولذلك فان :

العينة هى مجموعة كل المشاهدات عن خاصية معينة لظاهرة أو مجموعة من الظواهر التى لها خصائص مشتركة والتى تمثل جزءا من المجتمع .

وكما اوضحنا سلفا فانه بالرغم من أهمية الاحصاء الوصفى إلا أن المعلومات الاحصائية التى تنتج من العينة يستلزم تحليلها حتى يمكن تعميمها على المجتمع وهو ما يخرج عن نطاق البيانات . ولذلك فانه من الأهمية بمكان التحول من مجرد استخدام طرق لوصف البيانات الى طرق تفيد فى تعميم النتائج او بمعنى أخر التحول من الاحصاء الوصفى الى الاستنتاج الاحصائى

ولذلك فان الاحصائيين يوجهون معظم اهتماماتهم الى تحليل معلومات العينات الرقمية بدلا من مجرد جمع وعرض المعلومات أى بمعنى آخر يركزون اهتمامهم على التحليل الاحصائى أو ما يعرف بالاستنتاج الاحصائى أو الاحصاء التحليلى .

ولذلك فان :

الإحصاء التحليلى أو الإستنتاج الإحصائى عبارة عن ذلك الفرع من علم الإحصاء الذى يقدم ويستخدم الأدوات والوسائل الفنية للحصول على استنتاجات وتقديرات من خلال التحليل السليم للمعلومات الرقمية من العينات وتعميمها على المجتمع .

ومن ثم فان علم الاحصاء كما سبق تعريفه فى بداية هذا الباب يجمع بين الاحصاء الوصفى والاحصاء التحليلى إذ أن هذين الفرعين من علم الاحصاء يساعدان متخذو القرارات فى استخلاص النتائج من المعلومات غير الكاملة أو الكافية او المحدودة . ففى الاحصاء الوصفى فان الجداول والاشكال البيانية والمقاييس المختصرة تلخص البيانات وتعرض نمط التغير فى البيانات وفى المقابل نجد الاحصاء التحليلى يقدم الاستنتاجات والتقديرات المناسبة للأشياء غير المعروفة مع احتمالات الحصول على هذه التقديرات والاستنتاجات .

ولايضاح أهمية الاحصاء التحليلى فى تسهيل اتخاذ القرارات الرشيده نورد بعض الامثلة . فمثلا التنبؤ بالمبيعات لشركة معينة اعتمادا على المتغيرات المختلفة من انتاج ومخزون سلعى وأسعار وأموال سائلة أو نقدية وترويج واعلان الى غير ذلك مثل : التنبؤ بناتج الاستثمار فى مجموعة معينة من الاسهم والسندات اعتمادا على سلوك هذه الاسهم والسندات فى بورصة الأوراق المالية ومعدلات الربح للمشروعات وغير ذلك من المتغيرات والتنبؤ باستقصاء الرأى العام وفى الاختبارات التسويقية وانشاء نماذج الاقتصاد القياسى وكذلك التنبؤ بمعدل التضخم على المستوى القومى اعتمادا على قيمة العجز فى الموازنة العامة للدولة وسعر الفائدة الى غير ذلك من المتغيرات . وكمثال آخر التنبؤ بالاحتياجات المستقبلية للانشاءات كمشروعات الكهرباء والمياه والطرق والمدارس والمستشفيات اعتمادا على الامكانيات المادية المتاحة وعدد السكان الى غير ذلك من المتغيرات .

من كل ما تقدم يمكن القول أن الاحصاء التحليلى يلعب دورا هاما فى كل المجالات سواء على مستوى المشروع أو على المستوى القومى نظرا لزيادة حجم المشروعات وتعقدها وتعدد التكتلات الاقتصادية فى العالم والمتغيرات الدولية الأخرى من بروز نظام عالمى جديد تنفرد به الولايات المتحدة والتوجه الحالى نحو اقتصاد السوق وتغير مفهوم الوحدة الانتاجية من منظور داخلى بحت الى منظور كلى أو عالمى والتقدم التكنولوجى وثورة الاتصالات الرهيبة والتغيرات السياسية وشدة المنافسة فى الأسواق العالمية . ولذلك فان الاحصاء التحليلى يمكن المشروعات ان تحكم على عملياتها التجارية بالفشل أو النجاح

ودراسة السوق من حيث القوة الشرائية للمستهلكين وكذا عادات وتفضيلات المستهلكين والتنبؤ بالمبيعات ومراقبة جودة الانتاج والتنبؤ بالانتاج اعتمادا على تقديرات المبيعات للفترات المستقبلية ودراسة نظم الاستثمار الخاصة بالمشروعات وتحليل العمليات المالية لاتخاذ القرارات السليمة . وكذلك فانه على المستوى القومي يساعد الاحصاء التحليلي في التنبؤ بالناتج المحلى الاجمالى والدخل القومي والسكان والقوى العاملة ومستويات الأجور ومستوى المعيشة والايرادات السيادية من ضرائب وجمارك وغير ذلك من الظواهر المختلفة لاتخاذ القرارات المناسبة تحقيقا للأهداف الموضوعة .

يتضح من هذه الأمثلة السابقة ان هناك شيئا واحدا عاما بينها وهو عدم التأكد لأن هناك معلومات جزئية وغير كاملة أو غير مباشرة ولأن اتخاذ القرارات يتعلق بما سيحدث في المستقبل فانه يجرى استخدام الاحتمالات وكذلك هناك بدائل مختلفة يقوم الاحصائي بفحص وتقدير هذه البدائل المختلفة واختيار احدها لأن يكون الاختيار الأمثل أو الأحسن أو الأكثر مناسبه طبقا للظروف البيئية الداخلية والخارجية .

الا أنه للوثوق التام ان الاستنتاجات والتقديرت التى نتجت من التحليل السليم للمعلومات الرقمية من العينات صحيحة فان الأمر يستلزم قياس مأمونية هذه الاستنتاجات والتقديرات حيث ان المأمونية تمثل درجة خلو نتائج القياس من الخطأ أو درجة الاعتماد على الاستنتاجات والتقديرات لتعميمها على المجتمع

طبيعه البيانات المطلوبه للعمل الاحصائى

لاشك ان العمل الاحصائى يستلزم بيانات خام ومعلومات ذات طبيعة قاطعة ومناسبة ودقيقة ومطابقة وشاملة

وفى الحياة العملية نجد كثيرا من البيانات التى نحتاجها الإدارة فى التخطيط والتنسيق والمتابعة والرقابة والتنفيذ وحل المشاكل الإدارية فى المشروع هى تمثل بيانات داخلية.

ويقصد بالبيانات الداخلية البيانات المتعلقة بنشاط المشروع والتى يمكن الحصول عليها من دفاتر المشروع ومستنداته ، مثل البيانات عن انتاج المشروع ومبيعاته ومختلف العمليات الأخرى . ويمكن الحصول على هذه المعلومات من الحسابات الداخلية مثل حساب المشتريات وحساب أوراق القبض وحساب المبيعات ... الخ ويمكن وضع هذه الحسابات فى أنواع ثلاثة مثل التقارير المالية وتقارير العمليات وتقارير خاصة . ومن أهم التقارير المالية فى المشروع هى الميزانية العمومية التى تبين المركز المالى للمشروع فى وقت معين . وغالبا ما تلحق الميزانية بكشوف أو تقارير أخرى مثل تقرير عن التدفقات النقدية وتقرير عن أوراق القبض وتقرير عن المصاريف الرأسمالية والاستثمارات.

أما تقارير العمليات فيمكن تقسيمها تبعا للنشاطات المختلفة للمشروع أو تبعا للانتاج أو المنطقة . الخ . ومن أهم تقارير العمليات هى حساب الأرباح والخسائر (تقرير مالى) وكذلك تقرير عن المبيعات وعن الانتاج وتقرير عن المشتريات

وقد يحدث من وقت لآخر أن يحتاج المشروع الى تقارير اضافية لتحليل بعض البيانات التى لا تظهرها دفاتر المشروع . ومن ثم نجد تقارير منفصلة عن مشاكل معينة أو خاصة تحلل البيانات الداخلية بطريقة مختلفة عن الطريقة المحاسبية التى دونت بها البيانات .

كما قد يحتاج المشروع الى بيانات أخرى لمعرفة مركزه بالنسبة لنوع الصناعة التى يتبعها مثل مبيعات المشروع بالنسبة لمبيعات المشروعات الأخرى فى نفس الصناعة حتى يمكن عمل المقارنات المختلفة . وهذه البيانات هى فى الواقع عبارة عن بيانات خارجية من وجهة نظر المشروع . ويمكن تقسيم مصادر الحصول على بيانات خارجية الى مصادر تاريخية ومصادر الميدان . أما المصادر التاريخية فانها تشمل الوثائق والتقارير المنشورة عن الموضوعات المتعلقة بالبحث تحت الدراسة . وهناك كثيرا من الجهات الحكومية وغير الحكومية تقوم بنشر بيانات عن المجتمع . فمثلا توجد بيانات يقوم بنشرها الجهاز المركزى للتعبئة العامة والاحصاء وكذلك وزارات الاقتصاد والتخطيط وغيرها . ومن المستحسن استخدام البيانات التى تنشرها الجهات الأصلية بدلا من الالتجاء إلى الحصول على هذه البيانات من مصادر ثانوية . وذلك لان البيانات التى تنشرها مصادرها الرئيسية تتصف بعنصر التكامل عن البيانات المنشورة بواسطة المصادر الثانوية بالاضافة الى ذلك فان كثيرا ما نجد ان البيانات المنشورة بواسطة المصادر الأساسية تلحق بتقارير تصف الطرق المستخدمة فى جمع هذه البيانات وتفيد هذه التقارير الاضافية فى تقييم وتفسير البيانات المنشورة

وبجانب النشرات والاحصاءات التى يقوم بنشرها الجهاز المركزى للتعبئة العامة والاحصاء ، نجد ان هناك جهات أخرى تقوم بنشر بيانات مثل البنك المركزى والبنوك الكبرى بجمهورية مصر العربية واتحاد الصناعات والاتحاد العام للغرف التجارية واحصاءات الأمم المتحدة والأنترنت وغير ذلك.

وقد يعتبر بعض هذه النشرات مصادر أولية والبعض الآخر مصادر ثانوية . والمقصود بالمصادر الأولية الهيئة أو الهيئات التى قامت بجمع البيانات وتبويبها ثم نشرها مثل النشرات التى تصدر عن الجهاز المركزى للتعبئة العامة والاحصاء .

أما المصادر الثانوية فهى الهيئة أو الهيئات التى قامت بنشر البيانات رغم عدم قيامهم بجمع أو تبويب هذه البيانات . مثل قيام المؤلفين باقتباس بعض البيانات المطبوعة فى احدى نشرات الجهاز المركزى للتعبئة العامة والاحصاء فى كتاب أو مجلة أو بحث .

بعد ان اوضحنا المصادر التاريخية وما تتضمنها من مصادر أولية وثانوية فاننا نود أن نوضح ان قسما كبيرا من البيانات المعروضة فى المصادرالأولية قد سبق جمعها بواسطة البحوث الميدانية الاحصائية . ولذلك ففى الحالات التى يصعب فيها الحصول على البيانات الأساسية فى المصادر المنشورة تقوم الشركات والهيئات المختلفة باجراء دراسات للحصول على البيانات المطلوبة ومن ثم فان البحوث الميدانية هى المصدر الأصلى لكثير من البيانات المستخدمة بواسطة الباحثين الاحصائيين والاجتماعيين والإداريين .

وقد تجرى هذه البحوث الميدانية للحصول على معلومات عن الاشخاص ، رسالات المواد الأولية ، والعمليات التجارية وغيرها من الموضوعات المختلفة الا ان المشاكل الصعبة كثيرا ما تنشأ فى البحوث الميدانية المتعلقة بجمع بيانات عن الاشخاص ، ولذلك سنوجه اهتمامنا لهذا النوع من البحوث .

ومن الطبيعى أن الباحث سيجمع البيانات المطلوبة . أما بملاحظة الظاهرة بنفسه كما هو الحال فى التجارب المعملية أو أن يحصل عليها من مفردات مجتمع الدراسة الذين لهم اتصال بهذه الظاهرة كما هو الحال مثلا فى معرفة تفضيلات المستهلكين لسلعة معينة فى منطقة معينة ومن الواجب ان يصمم البحث الميدانى للحصول على بيانات دقيقة كافية للغرض من الدراسة وكذا الحصول على هذه البيانات بأقل تكلفة ممكنة .

ويتم جمع البيانات من الميدان من مفردات مجتمع أو عينة الدراسة الذين لهم اتصال بهذه الظاهرة عن طريق المقابلة الشخصية أو إرسال خطابات بالبريد أو المكالمات التليفونية أو الجمع بين بعض هذه الطرق أو كلها . ولا شك أن من الأدوات الرئيسية التى تستخدم مع مفردات مجتمع الدراسة فى الميدان هى استمارة الاستقصاء وهى الاستمارة التى تدون عليه البيانات من مفردات المجتمع تحت الدراسة وهى عادة ما تكون استمارة مطبوعة وهى الصحيفة التى يقوم بملئها مفردات المجتمع أو العينة تحت الدراسة بأنفسهم واعادتها للباحث.

وهناك أيضا نوع أخر من الاستمارات الاحصائية التى تستخدم فى جمع البيانات من المجتمع أو العينة تحت الدراسة هى كشف البحث وهى الصحيفة المطبوعة التى يقوم العدادون بملئها (وليس مفردات المجتمع أو العينة تحت

الدراسة) عن طريق المقابلة الشخصية أوالطرق الأخرى . ومن أهم الشروط الواجب توافرها عند تصميم استمارة الاستقصاء، تحديد التعاريف المستخدمة في الاستمارة تحديدا واضحا لاغموض فيه وان يكون عدد الاسئلة قليلا بقدر الإمكان وان تكون الاسئلة سهلة الألفاظ واضحة المعانى وأن تكون الاسئلة قصيرة وأن تكون اجابات الاسئلة قصيرة ويستحسن ان تكون اجابة السؤال بكلمة " نعم " أو " لا " . واذا كانت الاجابة عن سؤال معين يحتمل ان تكون له اجابات متعددة فيستحسن فى هذه الحالة كتابة كل الاجابات الممكنة الى جانب السؤال على أن يضع المجيب علامة صح أمام الاجابة المختارة كما يجب ان تغطى الاسئلة جميع أهداف البحث وعناصره الأساسية كما يجب ألا تكون الاسئلة محرجة أو تثير الاشمئزاز .

أنواع البيانات

طبقا لما توضح سلفا يلاحظ أن المشاهدات للمتغيرات النوعية تعرض فى صورة وصفية أى تعرض فى صورة كلمات مثل تقسيم الوحدات المنتجة فى مصنع معين الى وحدات جيدة ووحدات رديئة أو أن الهيكل الوظيفى فى شركة معينة عبارة عن كاتب ، مراجع ، مراقب ، رئيس قسم ، وكيل ، مدير وهكذا . الأنه للتعامل احصائيا مع هذه المشاهدات يجرى تجهيز البيانات فى صورة ارقام .أما المشاهدات للمتغيرات الكمية فهى متغيرات رقمية . الا أن الأرقام الموضوعة تختلف فيما بينها ، فمثلا عند وضع رقم كودى ٥ مثلا لاحدى الوظائف تختلف عن الرقم ٥ الذى يمثل حجم أو وزن الوحدة . ولذلك هناك أربعة أنواع للبيانات هى : البيانات الأسمية ، البيانات الترتيبية . بيانات الفترة ، بيانات النسبة .

(أ) البيانات الاسمية :

وهى عبارة عن مجرد ارقام تسمى أو تعنون، أو ترمز لايضاح الفروق فى البيانات النوعية (المتغيرات النوعية) بغرض تصنيف المشاهدات للمتغيرات النوعية . فمثلا فى استقصاء معين فانه يمكن ترميز الذين يجيبون بنعم بالرقم ١ والذين يجيبون بلا بالرقم صفر . كما يمكن ترميز الذين يجيبون بنعم بالرقم ١٠٠ والذين يجيبون بلا بالرقم ٢٠ . ويلاحظ أن هذه الأمثلة لاتمثل معنى معينا فى جمع أوطرح أو ضرب أو قسمة هذه الأرقام . ولذلك يستخدم فقط فى هذا النوع من البيانات أسلوب العد فقط ولايجوز استخدام خواص الجمع أو الطرح أو القسمة أو غير ذلك .

(ب) البيانات الترتيبية

وهى عبارة عن ارقام تحمل خواص البيانات الأسمية (للمتغيرات النوعية) بالاضافة الى ترتيب المشاهدات حسب الحجم اعتمادا على اهميتها . ويمكن استخدام هذه البيانات الترتيبيه فى اجراء المقارنات فى صورة "أكثر من" ، "أقل من" ، "يساوى" ، الا أن هذه البيانات لا تعطى حجم الفروق بينها أى بين بيانات أكثر من وبيانات أقل من وهكذا . فمثلا عند ترتيب الطلاب الناجحين حسب تقديراتهم ممتاز ، جيد جدا ، جيد ، مقبول يمكن ان تعرض كالآتى : ٤ ، ٣ ، ٢ ، ١ أو ٥٠٠ ، ٧٠ ، ٤٠ ، ٢٠ أو ١٠٠ ، ٢٠ ، ٥ ، ١ ويلاحظ أن النقطة الهامة هى أن الأرقام الكبيرة توضح التقييم الأفضل أو الترتيب العالى بينما الأرقام الأصغر توضح الترتيب الأدنى . ويلاحظ ايضا

(كما أوضحنا فى البيانات الترتيبية) أنه لايجوز أخذ الفروق بين هذه الأرقام لتمثل حجم التفضيل بل الأمر كله قاصر على الترتيب حسب الأهمية وليس حجم هذه الأهمية

(هـ) بيانات الفترة

وهى عبارة عن ارقام (للمتغيرات الكمية) تحمل خواص البيانات الترتيبية أى ارقام تمثل الفروق فى النوع والحجم بالاضافة الى ارتباط كل منها للاخر بفترات لان كل الأرقام يشار اليها بالنسبة لنقطة مفترضة تمثل الصفر . ولكن يلاحظ أنه يمكن اجراء الجمع والطرح وليس الضرب أو القسمة. ومن الأمثلة لذلك المشاهدات الخاصة بدرجات الحرارة .

(د) بيانات النسبة

ومن البيانات المقيسة أيضا ما يطلق عليها بيانات النسبة وهى عبارة عن ارقام تحمل خواص بيانات الفترة (أى أرقام لاتمثل فقط الفروق فى النوع عن طريق ترتيب المشاهدات حسب حجمها وتقسيمها باستخدام فترات) بالاضافة الى نسب ذات معنى لأن الأرقام يشار اليها بالنسبة نقطة مطلقة وهى الصفر والذى يوضح الغياب الكامل للخواص المقيسة ولذلك فان كل العمليات الحسابية من جمع وطرح وضرب وقسمة يمكن القيام بها . ولذلك فان بيانات النسبة تعتبر أعلى مستويات القياس . فمثلا عند قياس الأعمار ، الدخل ، الأجور ، الوزن الى غير ذلك تنتج بيانات النسبة ولذلك عند ترتيب عمر شخص يبلغ ٦٠ سنة أكبر من شخص عمر ٣٠ سنة بمقدار ٣٠ سنة فان ذلك يعتبر ترتيبا ذى معنى . كما يمكن القول ان عمر الشخص الأول يمثل مثلى الشخص الثانى الذى عمره ٣٠ سنة ولذلك فهى بيانات النسبة . ولذلك فان بيانات النسبة تعلو فى الأهمية بالنسبة لبيانات الفترة

الباب الثانى

اساليب عرض البيانات

اساليب عرض البيانات

تعتبر مرحلة جمع البيانات والمعلومات والحقائق عن المتغيرات والظواهر موضع الدراسة من أسس العمل الإحصائي التي لها أهمية خاصة لا يمكن إغفالها في أي دراسة علمية منظمة. وقبل الشروع في عملية جمع البيانات يجب أن يلم الباحث بعدة خطوات هامة وضرورة تميلها عليه طبيعة الدراسة يمكن أن نوجزها فيما يلي: -

أ - تحديد المشكلة العلمية أو تعيين مجال الظاهرة المراد دراستها وبحثها.
ب - الاتفاق على وحدة القياس التي ستستعمل في عملية جمع البيانات.
جـ - تعيين المتغيرات التي ستتناولها عملية القياس وحصر المصادر التي تعتمد عليها في الحصول على البيانات.
هـ - تحديد الأسلوب أو الطريقة التي تتبع في جمع البيانات والمعلومات.

وسوف نركز مناقشتنا في هذا الفصل حول الإطار العام لكيفية جمع البيانات من مصادرها المختلفة وما يتصف به كل مصدر من مزايا الاستخدام ومثالب ومشاكل التطبيق. وتجدر الإشارة هنا إلى أنه كلما كانت طريقة جمع البيانات سليمة كلما توفرت معلومات دقيقة عن مجموعة المتغيرات أو الظاهرة موضع الدراسة، وكلما أدى ذلك إلى رفع درجة الثقة في النتائج المستخلصة من التحليل الإحصائي، وبالتالي التوصل إلى قرارات سليمة غير متحيزة.

مصادر جمع البيانات Sources of Data

هناك مصدران أساسيان لجمع البيانات: الأول، يستمد منه الباحث المعلومات اللازمة لبحثة من بيانات تم جمعها وتجهيزها ونشرها بواسطة أجهزة متخصصة وأما المصدر الثاني فيعتمد فيه الباحث على نفسه في جمع وإعداد وتجهيز البيانات. ويعرف المصدر الأول بالمصدر غير المباشر بينما يطلق على المصدر الثاني «المصدر المباشر» أو «مصدر الميدان».

١ ـ المصدر غير المباشر في جمع البيانات

تتصف البيانات التي نحصل عليها من هذا المصدر بأنها بيانات غير أولية، تم تبويبها وتصنيفها من قبل بواسطة شخص آخر (غير الباحث) أو هيئة حكومية. ومن أمثلتها البيانات التي تتضمنها الدوريات والنشرات والكتب والتقارير والبحوث التي تصدرها وتنشرها الجهات والهيئات الحكومية ومراكز البحوث العلمية. ويلجأ الباحث إلى هذا المصدر في الحصول على البيانات التي يحتاج إليها بحثه في حالة وجود صعوبات (من حيث الوقت والتكاليف) تعترض عملية جمع البيانات من مصادرها الأولية، وعلى الرغم من سهولة وسرعة الحصول على البيانات من هذا المصدر، إلا أنه يعاب عليه صعوبة تحديد درجة الدقة أو الثقة في البيانات وعدم التأكد من سلامة الأعداد والتجهيز الإحصائي. لها وللتغلب على كل ذلك يجب على الباحث أن لا يتمادى في الاعتماد على هذا المصدر في حصوله على البيانات، وإذا كان مضطراً لذلك فيجب عليه الاعتماد على البيانات التي تصدرها أجهزة الإحصاء الرسمية في الدولة، مثل الجهاز المركزي للتعبئة العامة والإحصاء بجمهورية مصر العربية.

٢ ـ المصدر المباشر في جمع البيانات

تتميز البيانات التي يتم الحصول عليها من هذا المصدر بأنها بيانات أولية يعتمد الباحث في جمعها وتجهيزها للتحليل على نفسه. ويلجأ الباحث إلى هذا

المصدر في حالة إذا ما كانت طبيعة الدراسة تملى عليه الحصول على بيانات غير منشورة، أو نتائج بحوث سابقة تتعلق بموضوع البحث، كما في دراسة العلاقة بين العمليات البحرية Marine Processes الأمواج، التيارات. . . إلخ) والظاهرات التي تتأثر بها على ساحل منطقة ما في وقت معين. ومن مزايا المصدر المباشر في الحصول على المعلومات أن درجة الدقة وحدود الثقة في البيانات يمكن تحديدها عند تحليل البيانات كمياً، وهي في الغالب ما تكون مرتفعة مما يساعد بالتالي على استخلاص نتائج مرشوق فيها بدرجة كبيرة. إلا أن أهم المشاكل التي تواجه الاعتماد على المصدر المباشر هو الحاجة إلى الوقت والتكلفة المادية اللازمين لإنجاز مهمة الحصول على المعلومات. ونتيجة لذلك فإن الباحث يجد نفسه مضطراً إلى بذل قصارى جهده في جمع البيانات التي يحتاج إليها بالطريقة المباشرة في وقت بأقل تكلفة مادية ممكنة.

وعند جمع البيانات من مصادرها المباشرة فإن الباحث يعتمد على أحد الأسلوبين: أما أسلوب الحصر (المسح) الشامل وإذا لم يتيسر له جمع البيانات عن جميع مفردات المجتمع الأصلي فإنه يضطر إلى اختيار عينة، وهذا ما يطلق أسلوب المعاينة (العينات). ولكل من الأسلوبين جوانبه الإيجابية والسلبية التي نوضحها فيما يلي:

أولاً: أسلوب الحصر (المسح الشامل)

يعرف أسلوب الحصر الشامل أحياناً بأسلوب العد الكامل (أو التعداد Census) حيث أن معظم التعدادات تتم من خلاله، مثل التعداد السكاني Population Cencus والتعداد الزراعي أو التجاري أو الصناعي التي يعتمد عليها في استخراج بعض المقاييس والمؤشرات الإحصائية، والتي تكون أساساً في عملية التخطيط القومي أو وضع إطار عام للأبعاد الفعلية لإمكانية الدولة في مواجهة الأزمات الاقتصادية أو الاجتماعية وغيرها، والأساس في عملية جمع البيانات عن

طريق الحصر الشامل هو إدخال كل مفردات المجتمع الإحصائي، دون استبعاد أي مفردة، في البحث أو الاستقصاء. فمثلاً عند دراسة العمالة الصناعية في محافظة ما يقوم الباحث بعمل حصر شامل لجميع العمال حسب نوع كل صناعة، وكذلك عند دراسة التركيب المحصولي للأحواض الزراعية في أحد مراكز محافظة ما فإن الباحث يقوم بعمل حصر شامل لأنواع المحاصيل والمساحة التي تشغلها داخل كل حوض من الأحواض الزراعية. وبناء على ذلك فإن هذا الأسلوب يطبق عند دراسة المجتمعات الإحصائية مجهولة المعالم والتي تتطلب جمع بيانات شاملة عن كل مفردة من مفردات المجتمع حتى يمكن تحديد خصائصه ومعالمه بكل دقة وبدرجة عالية من الثقة.

ولأسلوب الحصر الشامل بعض المثالب والمشاكل عند استخدامه في جمع البيانات فهو لا يصلح للأبحاث التي يقترن استخلاص النتائج منها بوقت محدد، أو بمعنى آخر، أن هذا الأسلوب لا يتناسب مع الأبحاث التي يكون فيها لعنصري الوقت والتكاليف المالية أهمية خاصة وأثر كبير على استخلاص النتائج. وعلاوة على ذلك يتعرض تنفيذ أسلوب المسح الشامل في جمع البيانات كثير من الأخطاء التي من أهمها خطأ تحيز الباحث، سواء كان تحيز معتمد أو غير معتمد، الذي ينجم عن أخذ كل مفردات المجتمع في الدراسة حيث وجود احتمالات الخطأ في العد، أو الاحتمالات تجاهل بعض المفردات مما يؤثر على دقة النتائج. وللتخلص من خطأ هذا الأسلوب يمكن تقسيم المجتمع إلى مجموعات متجانسة لها خصائص متشابهة ومميزات مترادفة، ثم يجري البحث وعملية الحصر على كل قسم على حدة مع مراعاة التنسيق في الدراسة بين كل الأقسام. وأخيراً فإن الأسلوب يتطلب في إجرائه توفر جهاز فني إحصائي كبير واعتمادات مالية ضخمة ووقت متسع، مما يفسر أن معظم الدراسات والأبحاث التي تعتمد على هذا الأسلوب في إنجازها لا يقوم بها سوى أجهزة الإحصاء الحكومية مثل الجهاز المركزي للتعبئة العامة والإحصاء بجمهورية مصر.

ثانياً: أسلوب المعانية (العينات) Sampling

سبق أن عرفنا أن دراسة المجتمعات الإحصائية تعتمد أساساً على أخذ كل مفردات المجتمع للتعرف على خصائص ومعالم هذا المجتمع وبصفة عامة فإن معالم أي مجتمع (وهي مقادير ثابتة للمجتمع الواحد ولكنها تتغير من مجتمع إلى آخر) هي التي تعطي لهذا المجتمع صفاته دون غيره ونظراً لوجود صعوبات كثيرة تحول دون دراسة جميع مفردات المجتمع بواسطة أسلوب الحصر الشامل، فإننا نجري دراستنا على جزء صغير من هذا المجتمع أو ما يسمى بالعينة Sample وذلك اختصاراً للوقت وتوفير للجهد والنفقات، واتباع دراسة العينات أو أسلوب المعاينة يرفع مستوى العمل البحثي ويجعله أكثر دقة، وذلك لأن دراسة عدد قليل من المفردات أو الحالات يتيح للباحث فرصة جمع معلومات دقيقة وكثيرة عن كل مفردة أو حالة. وعلى العموم فإنه إذا ما وجدنا أنه من الضروري إجراء معاينة فإن رائدنا الأساسي يكون دائماً هو الحصول على عينة تعطي نتائجاً ذات دقة معينة بأقل تكاليف ممكنة. أو التي تعطي أعلى دقة بتكاليف محدودة.

ويفضل استخدام أسلوب المعاينة عند دراسة خصائص ومعالم المجتمعات اللانهائية مثل الوحدات الإنتاجية لإنتاج بعض الآلات، كما يفضل كذلك في الأبحاث العلمية التي تتطلب تصور عام أو رأي عام حول قضية أو مشكلة يراد دراستها في مجالات العلوم الطبيعية أو الاجتماعية. وفي كل من الحالات يجب أن تكون العينة ممثلة تماماً للمجتمع ولا تخضع للاختيار الشخصي. وذلك حتى يمكن الحصول بواسطة تطبيق الأساليب الكمية والمقاييس الإحصائية على نتائج يمكن تعميمها على المجتمع الأصلي المراد تحديد معالمه بدرجة عالية من الدقة والثقة. وتجدر الإشارة هنا إلى أنه عند دراسة العينات فإن المقاييس التي تحسب من توزيع العينة المختارة (مثل الوسط الحسابي والانحراف المعياري ـ سيأتي ذكرهما فيما بعد بالتفصيل) يسمى كل منها «إحصائية» وقيمة كل «إحصائية» تختلف من عينة إلى أخرى وللتفرقة بين المقاييس الیت نحسبها من العينة وتلك

التي نحصل عليها من دراسة جميع مفردات المجتمع بطريقة الحصر الشامل تسمى الأولى بالإحصائية Sample statistics بينما تعرف الثانية بالمعالم Parameters.

ويتوقف نجاح استخدام وتطبيق أسلوب المعاينة على عدة أمور هامة هي: تقدير حجم العينة، كيفية اختيار مفردات العينة من المجتمع، وتحديد نوع العينة، وفيما يلي مناقشة تفصيلية لكل منها على حدة:

(١) تقدير حجم العينة:

تتفق آراء كثير من الإحصائيين على أن حجم عينة البحث يتوقف على مجموعة من العوامل تنحصر في: الغرض من البحث حجم المجتمع الأصلي، مدى تباين الظواهر المختلفة في قطاعات المجتمع، درجة الدقة المطلوبة في البحث، البيانات المتاحة التي يمكن استخدامها في تعميم النتائج، والإمكانيات المادية. ونظراً لعدم وجود اتفاق بين الباحثين على وضع حد معين على أساس علمي - أو إحصائي - يحدد الحجم المناسب أو الأمثل للعينة لكي تمثل المجتمع الذي تسحب منه تمثيلاً جيداً، فإن تقدير حجم العينة - على مستوى معظم الدراسات والبحوث - تعتبر واحدة من المشكلات الخاصة بأسلوب المعاينة وتطبيق الأساليب الإحصائية، وفي مجال العمل الإحصائي يوجد اتجاهان عند تقدير حجم العينة:

الاتجاه الأول، يعتمد على الخبرة السابقة للباحث في هذا المجال، حيث أظهرت خلاصة الخبرات والتجارب أن حجم عينة في حدود ١٠٪ إلى ١٥٪ من حجم المجتمع الأصلي يبدو ملائماً في معظم الدراسات والبحوث. ويتميز هذا الاتجاه في تقدير حجم العينة بسهولته، كما أنه يفيد بعض الباحثين قليلي الخبرة في مجال العمل الإحصائي.

الاتجاه الثاني، يرتبط أساساً بنظرية الاحتمال Theory of probability مما يتطلب من الباحث الإلمام بقدر وافر من المعلومات الإحصائية والرياضيات حتى يستطيع استخدام الأساليب الإحصائية في تقدير الحجم الأمثل للعينة. ويعتمد هذا

الاتجاه على تحديد العوامل (المتغيرات) التي يتوقف عليها حجم العينة واعتبارها دلائل رئيسية أو مؤشرات أساسية لهذا الغرض وهو أمر يغفله الاتجاه الأول تماماً، كما يعتمد هذا الاتجاه على توفر بعض المعلومات عن حجم معالم المجتمع الأصلي عن طريق العينات التجريبية أو الاسترشادية Experimental or pilot sample . وتتمثل أهم العوامل والمتغيرات الرئيسية المحددة لحجم العينة في نسبة الخطأ المسموح به (أو درجة الدقة أو الثقة)، معامل التشتت (أو الانحراف المعياري) بين مفردات العينة أو المجتمع إن أمكن، والاختلاف النسبي بين المتوسط الحسابي للعينة ومتوسط المجتمع. وتوضع هذه المتغيرات في شكل صيغة رياضية تختلف باختلاف حجم العينة الاسترشادية، كما تترجم على هيئة معادلة خاصة في حالة إذا كان حجم المجتمع الأصلي الذي ستحيب منه العينة معلوماً.

فإذا ما تصورنا أن أحد الباحثين بصدد تقدير حجم عينة من مجتمع كبير غير محدود المفردات فإنه يقوم بسحب عينة استرشادية من هذا المجتمع وحساب بعض المقاييس الإحصائية منها لتقدير بعض خصائص أو معالم المجتمع، والتي عن طريقها يمكن تقدير حجم العينة المطلوب. فإذا كان حجم العينة الاسترشادية ٣٠ مفردة أو أكثر فإن أهم العوام المحددة لحجم العينة المطلوب تتمثل في:

أ ـ الانحراف المعياري بين مفردات العينة أو الخطأ المتوقع لمتوسط قيم مفردات العينة، وعنه يمكن تقدير الانحراف المعياري للمجتمع. أو ما يعرف بأحسن تقدير Best Estimate للانحراف المعياري بين مفردات المجتمع ويرمز له بالرمز $\hat{\text{ع}}$ ويحسب على أساس:

$$\hat{\text{ع}} = \text{ع} \times \sqrt{\frac{\text{ن}}{\text{ن} - ١}} \quad \text{أو} = \sqrt{\frac{\text{مجـ}(\text{س} - \bar{\text{س}})^{٢}}{\text{ن} - ١}}$$

$$\text{أو} \quad = \sqrt{\frac{\text{مجـ س}^{٢} - \text{ن}\,\bar{\text{س}}^{٢}}{\text{ن} - ١}}$$

حيث عـ هي الانحراف المعياري للعينة، س هي قيمة مفردة من مفردات العينة، س هي المتوسط الحسابي للعينة، ن هي الحجم الفعلي للعينة.

ب ـ خطأ المعاينة أو الخطأ المعياري Sampling of Stand and Error للمتوسط بين مفردات العينة أو مفردات المجتمع إن أمكن. وهو عبارة عن الدقة المطلوبة للتقدير الإحصائي من بيانات العينة، إذ أن تحديد حجم العينة يعتمد على الدرجة التي عندها يتجه متوسط العينة إلى الاختلاف والتباين عن متوسط المجتمع يرمز للخطأ المعياري بالرمز (خ م) ويحسب على أساس:

$$(خ م) = \frac{\hat{ع}}{\sqrt{ن}} \times \sqrt{١ - ف} \quad أو = \sqrt{\frac{(\hat{ع})^{٢}}{ن}} \times \sqrt{١ - ف}$$

$$أو \quad = \sqrt{\frac{(\hat{ع})^{٢}}{ن} \times (١ - ف)}$$

حيث ف تمثل نسبة حجم العينة ن إلى حجم المجتمع الأصلي ن (أي $\frac{ن}{ن}$)

وتسمى هذه النسبة «نسبة المعاينة Sampling Eraction أو معامل التصحيح لقيمة الخطأ المعياري للمجتمع الأصلي الذي يجب أن يكون أقل من خطأ المعاينة للمتوسط. وكلما كان زاد حجم العينة واقترب من حجم المجتمع الأصلي كلما اقتربت قيمة ف من الوحدة (الواحد الصحيح) وأصبحت قيمة معامل التصحيح صفراً وبالتالي فإن قيمة الخطأ المعياري تصير صفراً أيضاً.

جـ ـ القيمة المعيارية لاحتمال وقوع خطأ مسموح به ويرمز لها بالرمز (ز) ويمكن تحديد هذه القيمة من جدول التوزيع الاحتمالي الطبيعي إذا كان مستوى

الثقة Confidence Level الذي تعمم به النتائج على المجتمع معلوماً.

وإذا أخذنا في الاعتبار المتغيرات الثلاثة السابقة فإن حجم العينة يمكن أن يتحدد في ضوء تحديد الفارق الممكن التسامح فيه بين نتيجة العينة وما هو كائن فعلاً في المجتمع Tolerance (أي الخطأ المعياري) عند مستوى الثقة التي تعمم بها النتائج على المجتمع. ويمكن وضع هذا التصور لحجم العينة حيث أن:

$$(\text{خ م}) = \frac{\hat{\text{ع}}}{\sqrt{\text{ن}}}$$

$$\hat{\text{ع}} = \text{خ م} \times \sqrt{\text{ن}}$$

$$\sqrt{\text{ن}} = \frac{\hat{\text{ع}}}{\text{خ م}}$$

$$\text{ن} = \left(\frac{\hat{\text{ع}}}{\text{خ م}}\right)^{2} \quad \ldots\ldots\ldots\ldots\ (٢ ـ ١)$$

فإذا استطعنا أن نضع تقديراً مبدئياً للخطأ المعياري (خ م) الذي نرغب أن ننتهي إليه، وإذا استطعنا أيضاً سحب عينة استرشادية كبيرة (من الثابت إحصائياً أنه إذا بلغ حجم العينة ٣٠ مفردة أو أكثر فإنه يمكن أن يعطي حدوداً مرتفعة من الثقة لتقدير متوسط المجتمع، وانحرافه المعياري من متوسط العينة وانحرافها المعياري ويرجع ذلك إلى أنه كلما حجم العينة كثيراً كلما أدى ذلك إلى تكوين توزيع طبيعي)، فإن باستطاعتنا تقدير الحجم الأمثل بالاستعانة بالمعادلة السابقة (٢ ـ ١).

مثـال (١):

على أساس عينة استرشادية تتكون من ١٠٠٠٠ عامل قدر أن متوسط إنتاج العامـل من الملابس على مستوى الجمهوريـة هو ٥ وحدات وأحسن تقديـر للانحراف المعياري ($\hat{ع}$) لإنتاج الملابس على المستوى القومي ٢ وحدة، وأن خطأ المعيـاري للمتوسط هـو ٠٢ر٠ وحدة. فلو افترضنـا أننا نريد تقديـر المتوسـط القومي لإنتاج العامل من الملابس لأقرب ـــ وحدة عند مستوى احتمالي ٦٨ر٠ (أي بمستوى الثقة ٦٨٪) فإن أقل حجم مطلوب للعينة في هذه الحالة يكون:

$$\text{حجم العينة (ن)} = \left(\frac{\hat{ع}}{خ\,م}\right)^{٢}$$

$$= \left(\frac{٢}{٢٥ر}\right)^{٢} = (٨)^{٢} = ٦٤ \text{ مفردة}$$

وعلى ذلك فإن أي عينة مكونة من ٦٤ مفردة تكون كافية لإعطاء تقدير للمتوسط القومي (أي متوسط المجتمع) بدقة $\pm \frac{١}{٤}$ وحدة وبمستوى ثقة ٦٨٪.

ولو افترضنا ـ مرة أخرى ـ أن درجة الدقة المطلوبة لتقدير المتوسط القومي لإنتاج العامل من الملابس هي نفس الدقة السابقة (أي إلى أقرب $\frac{١}{٤}$ وحدة) ولكن عند المستوى الاحتمالي ٩٥ر٠ (أي بمستوى الثقة ٩٥٪) الـذي تكون عنده حدود الثقـة عبارة عن خطـأين معيارين للمتوسط (أي $\pm$ ٢ × خ م). ومعنى ذلك أن ٢ × قيمة الخطأ المعياري للمتوسط لا بد أن تساوي الدقة المطلوبة لحساب المتوسـط العام وهي ٢٥ر٠ وحـدة، وبعبارة أخرى فإن الخطأ المعياري

للمتوسط، لدرجة ثقة ٩٥٪، يساوي $\frac{٠٫٢٥}{٢}$ أي ١٢٥ر وحدة وبتطبيق المعادلة (٩ ـ ١) فإن حجم العينة يكون في هذه الحالة.

$$\text{حجم العينة (ن)} = \left(\frac{\hat{ع}}{خ\ م}\right)^{٢}$$

$$= \left(\frac{٢}{١٢٥ر}\right)^{٢} = (١٦)^{٢} = ٢٥٦ \text{ مفردة}$$

وبناء على ذلك فأننا لكي نحصل على تقدير للمتوسط القومي لإنتاج العامل من القمح لأقرب $\frac{١}{٤}$ وحدة بمستوى ثقة حجم العينة اللازم لا بد أن يكون ٢٥٦ مفردة.

وهناك صيغة أخرى لتحديد الحجم الأمثل للعينة تأخذ في اعتبارها المتغيرات السابق ذكرها والتي تحسب من عينة استرشادية يبلغ حجمها ٣٠ مفرد أو أكثر وهذه الصيغة هي:

حجم العينة =

$$\frac{\text{الانحراف المعياري للعينة} \times \text{القيمة المعيارية لاحتمال خطأ مسموح به درجة معينة}}{\text{الدقة المطلوبة للتقدير الإحصاري أو الخطأ المعياري}}$$

أي أن:

$$ن = \left(\frac{عـ \times ز}{خ\ م}\right)^{٢} \quad .... \quad (٢ ـ ٢)$$

مثال (٢):

إذا كان الانحراف المعياري لعينة استرشادية مكونة من ٣٠ عاملاً لدراسة مستوى المعيشة لمجتمع عمالي هو ١٠ جنيهات شهرياً وأن الخطأ المعياري المسموح به لتقدير المتوسط العام للدخل الشهري هو ٥ر٢ جنيهاً وذلك بمستوى

ثقة ٩٥٪، فإن الحجم الأمثل للعينة الذي يحقق الدقة المطلوبة يمكن تقديره بعد تحديد قيمة (ز) المعيارية من جدول التوزيع الطبيعي المناظرة لمستوى الثقة ٩٥٪ وهي في هذه الحالة تساوي ٢ تقريباً. وعلى ذلك فإن:

$$\text{حجم العينة (ن)} = \left(\frac{١٠ \times ٢}{٢{,}٥}\right)^{٢} = (٨)^{٢}$$

= ٦٤ عـــامـــلاً

وبما أن حجم العينة الاسترشادية ٣٠ عاملاً فإننا نحتاج إلى ٣٤ عاملاً آخر ليصبح الحجم الفعلي للعينة ٦٤ عاملاً، والذي فيه يمكن تقدير المتوسط العام لدخول المجتمع العمالي قيد البحث بالدقة المطلوبة (أو الخطأ المعياري للمتوسط) المشار إليها.

ويمكن أيضاً تحديد حجم العينة التي تحقق الدقة المطلوبة أو الخطأ المسموح به في حساب المتوسط العام للمجتمع من إحصائية عينة تجريبية صغيرة يقل عدد مفرداتها عن ٣٠ مفردة وفي هذه الحالة تؤخذ العوامل السابقة المحددة لحجم العينة في الاعتبار مع اختلاف واحد وهو أن القيمة المعيارية المناظرة لاحتمال وقوع خطأ مسموح به بدرجة معينة في جدول التوزيع الطبيعي يجب أن تستبدل بقيمة معيارية أخرى من جدول توزيع «ستيودنت ـ ت» مناظرة لعدد من المفردات يقل عن مفردات العينة التجريبية بمفردة واحدة، وعند مستوى الدلالة أو الثقة المطلوب.

وبناء على ذلك فإن:

$$\text{حجم العينة} = \frac{\text{مربع (الانحراف المعياري للعينة} \times \text{قيمة ت المعيارية)}}{\text{مربع الدقة المطلوبة أو الخطأ المعياري}}$$

أي أن

$$ن = \frac{(د \times ت)^{2}}{(خ م)^{2}} \quad \ldots \ldots \ldots \ldots \ldots \quad (٢ - ٣)$$

مثـال (٣)

في دراسة اجتماعية عن مستوى المعيشة لعمال إحدى الصناعات التي يبلغ عدد مصانعها ٩٢ مصنعاً، يراد معرفة مصانعها ٩٢ مصنعاً، يراد معرفة حجم عينة المصانع يمكن منه تقدير متوسط الدخل السنوي لجميع العمال وذلك في حدود ٥٠ جنيهاً زيادة أن نقصان عن متوسط دخل عمال مصانع العينة بمستوى ثقته ٩٥٪، (أي بمستوى دلالة ٥٪)، فإننا في هذه الحالة نأخذ عينة تجريبية مكونة من دخول عمال ٢٥ مصنعاً ونحسب منها متوسط الدخل والانحراف العياريد ونفترض أنه كان ٣٣٢ جنيهاً و ٣٣٢٫٥ جنيهاً و ١٦١٫٥ جنيهاً على الترتيب. وبما أن عدد مفردات العينة التجريبية أقل من ٣٠ مفردة، فلا بد إذن من تعيين قيمة ت من جدول «توزيع» «استيودنت ـ ت» المناظرة لعدد مفردات يقل عن مفردات العينة بمفردة واحدة أي (٢٥ ـ ١)، وعند نسبة الخطأ المسموح بها وهي ٥٪، وباستخدام هذه المؤثرات فإن قيمة «ت» تساوي ٢٫٠٦٠ وبذلك فإننا يمكن أن نطبق المعادلة (٢ ـ ٣) لنحصل على حجم العينة المطلوب كما يلي: -

$$ن = \frac{(١٦١٫٥ \times ٢٫٠٦٠)^{2}}{(٥٠)^{2}} = ٤٤٫٢ \text{ مفردة}$$

وبناء على ذلك فإن حجم عينة مكونة من عمال ٤٥ مصنعاً تكون كافية لإعطاء صورة صادقة عن الدخل المستوى لعمال جميع المصانع. ومن المعادلة السابقة (٢ ـ ٣) يمكن أن نستنتج أنه كلما كبرت قيمة الانحراف المعياري للعينة التجريبية كلما زاد حجم العينة المطلوب والعكس يحدث مع الدقة المطلوبة أو

الخطأ المعياري لحساب المتوسط العام، أي أنه كلما قلت أم انخفضت هذه الدقة أو زادت قيمة الخطأ المعياري كلما قل حجم العينة ففي المثال السابق إذا انخفضت الدقة في حساب المتوسط العام ليصل إلى ١٠ جنيهاً زيادة أو نقصان عن متوسط العينة، فإن حجم العينة المطلوب سيكون ١٢ مصنعاً فقط، وهو حجم كاف أيضاً لإعطاء صورة عامة عن متوسط الدخل السنوي لعمال جميع المصانع بالدقة المطلوبة

كما يمكن تحديد حجم العينة عن طريق تحديد النسبة المئوية لوجود الظاهرة موضع الدراسة في العينة التجريبية وتعيين مستوى الثقة التي تعمم بها النتائج على المجتمع وتحديد الفارق الممكن التسامح فيه بين نتيجة العينة وما هو كائن فعلاً في المجتمع، وهذا يتطلب تحديد الخطأ المعياري للعينة على النحو التالي: -

$$\text{الخطأ المعياري (خ م)} = \sqrt{\frac{\text{أ} \times \text{ب}}{\text{ن}}} \quad .... \quad (٢ - ٤)$$

حيث أ هي النسبة المئوية لوجود الظاهرة، ب هي النسبة المئوية لعدم وجود الظاهرة (أي ١٠٠ مطروحاً منها نسبة وجود الظاهرة)، ن هي حجم العينة. ويمكن نقل هذه المعادلة على النحو التالي:

$$\text{ن} = \frac{\text{أ} \times \text{ب}}{(\text{خ م})^{2}} \quad ... \quad ... \quad (٢ - ٥)$$

وفي كل الحالات إذا استطعنا أن نضع تقديراً مبدئياً لدقة القياس أو الخطأ المعياري الذي ترغب أن ننتهي إليه. إذا استطعنا كذلك تقدير نسبة وجود الظاهرة في الحالات المدروسة. فإنه باستطاعتنا تقدير الحجم الأمثل للعينة بالمعادلة السابقة (٢ - ٥)

مثـال (٤):

من عينة استرشادية لدراسة مدى تأثير البرامج التلفزيونية على ثقافة سكان أحد الأقسام الإدارية بمحافظة ما وجد أن النسبة المئوية لحائزي الأجهزة التلفزيونية هي ٨٦ ٪ بدقة تصل إلى ١٠ ٪، والمطلوب تحديد الحجم الأمثل للعينة التي يمكن عن طريقها دراسة هذا التأثير بمستوى ثقة ٤ر٩٥ ٪ فبما أن مستوى الثقة التي تعم بها النتائج على المجتمع هو ٩٥ ٪ فإن حدود هذه الثقة Confidence linits عبارة عن ± خطأين معيارين (أي ± ٢ × خ م) وقيمتيهما لا بد أن تساوي ١٠ ٪ومعنى ذلك أن الخطأ المعياري في هذه الحالة هو ٥٪ (أي $\frac{١٠}{٢}$٪) وأن حجم العينة المطلوب حسب المعادلة (٢ ـ ٥) هو:

$$\text{حجم العينة (ن)} = \frac{\text{أ} \times \text{ب}}{(\text{خ م})^{٢}}$$

$$= \frac{٨٦ \times ١٤}{(٥)^{٢}}$$

$$= \frac{١٢٠٤}{٢٥} = \text{٢ر٤٨ مفردة}$$

وعلى ذلك فإن الحجم الأمثل للعينة والذي نأمل أن يحقق الدقة المطلوبة في هذا المثال هو ٤٩ حائزاً لأجهزة التلفزيون في المنطقة موضع الدراسة.

وتستخدم صيغة أخرى لتحديد الحجم المناسب للعينة اعتماداً على خصائص بيانات عينة استرشادية يمكن منها تعيين النسبة المئوية لوجود الظاهرة موضع الدراسة بالإضافة إلى تقدير الخطأ المعياري وتحديد مستوى الثقة التي تعمم النتائج هذه الصيغة هي:

$$ن = أ \times ب \times \left(\frac{ز}{خ م}\right)^2 \qquad (٢ - ٦)$$

حيث ز هي القيمة المعيارية لاحتمال وقوع خطأ مسموح به عند مستوى معين.

مثال (٥):

من عينة تجريبية تتكون من ٣٠ ناخباً وجد أن ١٢ ناخباً سيقومون بإعطاء أصواتهم المرشح الحزب «أ»، فإذا أريد تقدير نسبة الناخبين الذين سيدلوا بأصواتهم لانتخاب مرشح هذا الجزب من جملة الناخبين بدقة (أو خطأ معياري) ± ١٪ وبمستوى ثقة ٩٥٫٤٪ فإن حجم العينة المطلوب لتحقيق ذلك يتحدد على أساس:

أ - نسبة الناخبين في العينة = $\frac{١٢ \times ١٠٠}{٣٠}$ = ٤٠٪

ب - القيمة المعيارية (ز) لمستوى الثقة ٩٥٪ = ٢ تقريباً.

جـ - الخطأ المعياري للتقدير = ١٪.

وعلى ذلك فإن:

حجم العينة (ن) = $٤٠ \times ٦٠ \times \left(\frac{٢}{١}\right)^2$

= ٢٤٠٠ × ٤ = ٩٦٠٠ مفردة أو ناخباً

وتجدر الإشارة إلى هنا إلى أن معظم استطلاعات الرأي Opinion Polls في النواحي السياسية تعتمد على عينات يصل حجم أي منها إلى ٢٠٠٠ مفردة تقريباً، حتى يكون الخطأ المعياري المتسامح فيه بين نتيجة العينة وما هو كائن فعلاً في

المجتمع ± ٢٪ أو أكثر قليلاً، وذلك قبل أن يدخل في الاعتبار عوامل أخرى مثل خطأ التحيز أو تغيير الرأي في آخر دقيقة تجاه موضوع الاستطلاع مثل ترشيح عضو أحد الأحزاب من جانب نسبة من المبحوثين أو الناخبين.

كما يمكن تحديد حجم العينة عن طريق معرفة حجم المجتمع الأصلي فقط وذلك عن طريق تحديد مجموعة من العوامل أو المحددات الرئيسية التي يمكن أن نجملها فيما يلي: -

أ - حجم المجتمع الأصلي الذي ستسحب منه العينة ويرمز له بالرمز ن.

ب - معامل التشتت بين مفردات العينة ويرمز له بالرمز (م) ويحسب على أساس:

$$(م) = \frac{\text{الانحراف المعياري}}{\text{المتوسط الحسابي}} \times ١٠٠$$

جـ - مربع متوسط معامل التشتت للمتوسط بين مفردات العينة ويرمز له بالرمز (م س)^2، ويحسب على أساس:

$$(م_{س})^2 = (\frac{م^2}{ن} \times ١ - ف) \quad .... \quad (٢ - ٧)$$

حيث هي حجم العينة، ف هي نسبة المعاينة أي نسبة حجم النسبة إلى حجم جتمع الأصلي.

د - الفارق النسبي بيا المتوسط الحسابي للعينة ومتوسط المجتمع ويرمز له بالرمز (ق)، ويمكن حسابه كما يلي: -

الفارق النسبي = الجزر التربيعي لعامل التشتت بين مفردات العينة × القيمة المعيارية للدقة المطلوبة بدرجة معينة.

أي أن:

ق ـ (م$_{س}$ × ز) .... (٢ ـ ٨)

وبناء على العلاقات الرياضية المتبادلة بين هذه العوامل الأربعة وحجم العينة فإننا يمكن أن نحدد حجم العينة المطلوب في ضوء المعادلة الآتية:

$$\text{حجم العينة} = \frac{\text{حجم المجتمع الأصلي}\times\text{مربع القيمة المعيارية}\times\text{مربع معامل التشتت}}{\text{حجم المجتمع الأصلي}\times\text{مربع الفارق النسبي}+\text{مربع القيمة المعيارية}\times\text{مربع معامل التشتت}}$$

أي أن:

$$\text{ن} = \frac{\text{ن} \times (\text{ز})^{2} \times (\text{م})^{2}}{\text{ن} \times (\text{ق})^{2} + (\text{ز})^{2} \times (\text{م})^{2}} \quad (٢ ـ ٩)$$

مثـال (٦):

يريد أحد الباحثين تحديد حجم عينة من مجتمع يحتوي على ٢٠٠٠ مفرد وذلك في ضوء الافتراضات الآتية التي يريد أنها ضرورية لتطبيق الطرق الإحصائية واستخلاص النتائج التي على أساسها تتخذ القرارات اللازمة: ـ

أ ـ معامل التشتت بين مفردات العينة (م) في حدود ٣٠٪.

ب ـ نسبة الخطأ المسموح به لا تزيد عن ٥٪ أي أن تعمم النتائج بثقة قدرها ٩٥٪.

جـ ـ بما أن القيمة المعيارية المناظرة لاحتمال وقوع الخطأ المسموح به وهو ٥٪ في جدول التوزيع الطبيعي (الاعتدالي) تساوي ٢ تقريباً فإن: الفارق النسبي (ن) = ٢٠ر٠×٢= ٠٤ر٠.

وعلى ذلك فإن حجم العينة المطلوب يمكن تحديده كما يلي:

$$\text{حجم العينة (ن)} = \frac{١٠٠ \times (٢)^{2} \times (٣ر٠)^{2}}{٢٠٠٠ \times (٠٤ر٠)^{2} + (٢)^{2} \times (٢ر٠)^{2}}$$

$$= \frac{٣٢٠}{٣٫٣٦} = ٩٥٫٢٥$$

وعلى ذلك فإن حجم العينة المناسب والذي بحفظ افتراضات الباحث هو ٩٦ مفردة تقريباً من مجتمع يحتوي على ٢٠٠٠ مفردة.

وهناك طريقة أخرى لتحديد حجم العينة يمكن استخدامها إذا كان حجم المجتمع الأصغر معروفاً، وذلك بعد تحديد الحجم التقريبي للعينة والذي يتطلب:

أ ـ تحديد الدقة المطلوبة أو الخطأ الذي يمكن التسامح بين نتيجة العينة وما هو كائن فعلاً في المجتمع.

ب ـ تحديد مستوى الثقة التي تعمم بها النتائج على المجتمع

جـ ـ اختيار النسبة المئوية لوجود الظواهر قيد البحث (ح) التي تحقق أكبر رقم إذا ما ضربت في النسبة المئوية المكملة (١٠٠ ـ ح).

وتتخذ معادلة هذه الطريقة الشكل الآتي:

$$ن_1 = \left[\frac{ح \times (١٠٠\% - ح)}{(خ)^2} \right] \times (ز)^2 \quad \ldots\ldots \quad (٢ - ١٠)$$

حيث ح هي نسبة وجود الظواهر قيد البحث وتمثل ٥٠٪ $ن_1$ هي الحجم التقريبي للعينة.

وبعد الحصول على الحجم التقريبي للعينة، يتعين تحديد الحجم الفعلي لها في ضوء حجم المجتمع الأصلي (ن) وذلك باستخدام معادلة التصحيح وهي:

$$\text{الحجم الفعلي للعينة (ن)} = \frac{ن_1}{١ - \frac{ن_1 \cdot ١}{ه}} \quad (٢ - ١١)$$

مثـال (٧): ـ

لنفرض أن أحد الباحثين يريد تحديد حجم عينة من مجتمع يحتوي على ٤٠٠٠ مفردة بناء على بعض الافتراضات التي رأها ضرورية في هذا الصدد هذه الافتراضات هي. ـ

أ ـ نسبة الخطأ المسموح به (أو الدقة) في حدود ± ٥٪

ب ـ مستوى الثقة التي تعم بها النتائج لا تقل عن ٩٥ ٪.

جـ ـ نسبة وجود الظواهر موضع البحث في العينة ٥٠ ٪ ونسبة عدم وجودها ٥٠ ٪ أيضاً.

وباستخدام هذه الافتراضات والتي يمكن أن تتحقق من تحديد حجم مناسب للعينة، نجد أن: عند مستوى الثقة ٩٥ ٪ تكون القيمة المعيارية المناظرة في جدول التوزيع الطبيعي هي ٢ تقريباً. وعلى ذلك فإن حجم العينة المطلوب يتحقق بتطبيق المعادلتين (٢ ـ ١٠) أو (٢ ـ ١١) كما يلي:

$$\text{الحجم التقريبي للعينة} = \frac{\text{٥٠} \times \text{٥٠}}{(\text{٥})^{\text{٢}}} \times (\text{٢})^{\text{٢}}$$

$$\text{الحجم التقريبي للعينة} = \frac{\text{٤} \times \text{٢٥٠٠}}{\text{٢٥}} = \text{٤٠٠ مفـــــردة}$$

$$\text{الحجم الفعلي للعينة} = \frac{\text{٤٠٠}}{\text{١} - \dfrac{\text{٤٠٠} - \text{١}}{\text{٤٠٠٠٠}}}$$

$$= \frac{\text{٤٠٠}}{\text{٠ر٩٠٠٢٥}} = \text{٤٤٤ر٣٢ مفـــردة}$$

وعلى ذلك فإن حجم العينة المناسب والذي يمكن أن يحقق افتراضات

الباحث هو ٤٤٥ مفردة من مجتمع يحتوي على ٤٠٠٠ مفردة أي نسبة ٩ ٪ تقريباً من حجم المجتمع الأصلي.

أما إذا تجمعت لدينا بيانات خاصة من معالم المجتمع الأصلي (المتوسط الحسابي والانحراف المعياري) الذي ستسحب منه العينة دون أن نتمكن من معرفة حجمه، فإننا نستخدم طريقة المنحنى الطبيعي (الاعتدالي) للحصول على الحجم المناسب أو الأمثل للعينة. ويوضح ذلك المثال التطبيقي الآتي:

مثال (٨):

إذا كان متوسط الإنتاج القومي للقمح في عام ما هو ٢ر٧ أردب ٪ فدان والانحراف المعياري لهذا المتوسط هو ٥ر١ أردب، فما هو حجم العينة (فدان) التي ينبغي اختيارها من محافظة ما بشرط ألا يكون خطأ الصدفة أكثر من ٥٠ ٪ وأن يكون متوسط الإنتاج في (العينة ٠٧ر٧ أردب)؟

الحل: -

نظراً لأننا افتراضنا أن نسبة خطأ الصدفة هي ٥٪، فإنه يمكن إيجاد القيمة المعيارية (ز) والتي تعرف بالمتغير المعياري Standard Variable من الجداول الإحصائية الخاصة بتحديد المساحة تحت المنحنى الطبيعي (الاعتدالي)[1]. ثم تطبق المعادلة الآتية: -

$$\text{المتغير المعياري} = \frac{\text{الفرق بين متوسطي المجتمع والعينة} \times \sqrt{\text{حجم العينة}}}{\text{الانحراف المعياري لمتوسط المجتمع}}$$

(١) انظر جدول تحديد المساحة تحت المنحنى الطبيعي (الاعتدالي) ضمن ملاحق هذا الكتاب

أي أن

$$(ز) = \frac{م - \bar{س} \times \sqrt{ن}}{ع} \qquad (٢ - ١٢)$$

وعلى ذلك فإن:

$$١٫٦٤ = \frac{٢٫٧ - ٠٫٧ \times \sqrt{ن}}{١٫٥}$$

$$٢٫ \sqrt{ن} = ٢٫٤٦$$

$$\sqrt{ن} = \frac{٢٫٤٦}{٢٫} = ١٢٫٣$$

ن = ١٥١٫٢٩ فداناً

بمعنى أنه ينبغي أن تسحب عينة حجمها ١٥٢ فداناً من المحافظة قيد البحث لكي تحقق الدقة أو الخطأ الصدقة والمتوسط المطلق للعينة.

من كل مما سبق يمكن القول أن حجم العينة الذي نحصل عليه بإحدى المعادلات السابقة لا يعتبر ملزماً الآن الافتراضات التي تقوم عليها هذه المحاولات غير ملزمة لأي دراسة، وكل ذلك ما هو إلا مجرد علامات تحدد أسلوب العمل في هذا المجال في حدود أقل خطأ ممكن بطريقة موضوعية غير متحيزة.

(٢) اختيار مفردات العينة

بعد أن تعرفنا على الطرق المختلفة التي تحدد الحجم المناسب للعينة التي سيجري عليها البحث والاستقصاء، فإننا الآن بصدد التعرف على طريقة (عملية) اختيار مفردات هذه العينة من بين مفردات المجتمع الأصلي، أو ما يعرف بأسلوب سحب العينة من المجتمع وعملية اختيار مفردات العينة كواحدة من المشكلات الخاصة بأسلوب المعاينة، تتوقف أساساً على خجم المجتمع الأصلي. فإذا كان

حجم المجتمع صغيراً أي مشتملاً على عدد محدد Finite من المفردات فإن المشكلة لا تكون مشكلة اختيار العينة من بين مفردات المجتمع، بل تكون مشكلة الحصول على عدد كاف من المفردات لغرض البحث. فمثلاً إذا أراد الباحث أن يجري دراسة على كبار الزراعين بإحدى القرى كنموذج لنفس الفئة في القطر فقد يحدد هذه الفئة بأنها تشتمل على كل من يمتلك ١٠٠ فداناً أو أكثر من الأراضي الزراعية في القرية. وفي هذه الحالة قد يكون عدد هؤلاء الملاك قليلاً لدرجة أن العينة تستنفذهم جميعاً. كما تكون عملية الاختيار من المجتمع الأصلي عملية مشترطة بشروط تحدد المفردات (عدد الملاك) التي تكون منها العينة المطلوبة. وبالطبع كلما كثرت الشروط اللازمة للعينة كما صعب الحصول عليها وكلما قل عدد المفردات الذين يتم الاختيار من بينهم. أما إذا كان حجم المجتمع الأصلي كبيراً جداً أي مشتملاً على عدد غير محدود من المفردات المستوفية لجميع الشروط اللازمة في العينة فإنه من اللازم إجراء عملية اختيار مفردات العينة أما بواسطة الاختيار غير العشوائي (المعاينة العمدية) أو بواسطة الاختيار العشوائي. وقبل أن نوضح كيفية الاختيار في كل من الطريقتين، فإنه يجدر بنا أن نذكر الشروط التي ينبغي توافرها في العينة حتى نستخلص بها عن المجتمع الأصلي الكبير.

سبق أن قلنا أنه يجب أن تكون العينة ممثلة للمجتمع أو بمعنى آخر يجب مراعاة أن تمثل العينة جميع أفرد المجتمع، وألا تكون متحيزة bias لجزء أو أجزاء من المجتمع الأصلي لأنه يتوقف على العينة المنتقاة كل قياس أو نتيجة يخرج بها الباحث. فمثلاً إذا أردنا سحب عينة لتقدير متوسط الدخل من الإنتاج الزراعي لملاك الأراضي الزراعية على مستوى أحد المراكز، فإذا أخذت عينة من فئة الملاك الذين يملكون ١٠٠ فداناً وأكثر فإن العينة تكون غير ممثلة لمجتمع الملاك حيث أن هذه الفئة تمثل نسبة صغيرة جداً من جميع الملاك، وبالتالي لا بد أن تحتوي العينة على ملاك من جميع فئات الملكية وبناء على ذلك يجب أن تتوفر في العينة الممثلة Representative Sample مجموعة من الشروط يمكن تلخيصها في شرطين أساسيين هما: ـ

١ ـ تكون مفردات العينة ممثلة للمجتمع الذي يجري عليه البحث تمثيلاً صحيحاً وليست ممثلة لمجتمع آخر. بمعنى أنه إذا تكررت نفس النتائج على عينات، أخرى من نفس المجتمع كانت العينة التي يجري عليها البحث عينة ممثلة للمجتمع الأصلي أصدق تمثيل، وبذلك يمكن أن تكون خصائص مفردات العينة (إحصائياً العينة) متقاربة أو متشابهة مع خصائص المجتمع (معالم المجتمع) الذي تنتمي إليه.

٢ ـ ألا تكون المفردات المختارة ممثلة لجزء (قطاع) من أجزاء المجتمع الأصلي بل يجب أن تمثل جميع أجزاء المجتمع، وهذا يتطلب تكوين إطار المعاينة الذي تؤخذ منه العينة.

وإطار المعاينة Samling Frame وهو المصدر الذي تؤخذ منه العينة أو بعبارة أخرى هو حصر شامل (القائمة أو الدليل) لجميع المفردات وحدات المجتمع الأصلي المراد دراسته مثال ذلك قائمة بأسماء العمال في أحد المصانع، أو مختلف أنواع الرواسب التي توجد على الشاطىء أو مواقع المحلات العمرانية الريفية على خريطة إحدى الدول وعند اختيار العينة من المجتمعات المحدودة يقسم المجتمع الأصلي للظاهرة قيد البحث إلى عدة أقسام تسمى وحدات لمعاينة Sampling Units (شخص، أسرة، قرية) ويكون إطار العينة حينئذ عبارة عن القائمة أو مجموعة القوائم التي تتضمن الوحدات التي تتألف منها المجتمع. ويشترط في إطار المعاينة أن يكون شاملاً لجميع مفردات المجتمع التي يمكن الوصول إليها بسهولة وذلك حتى يكون اختيار العينة سليماً. كما يشترط أن يكون إطار المعاينة متجدداً حتى تعطي المفردات أو الوحدات التي تستجد على الإطار القديم نفس الفرصة في الظهور.

وكما ذكرنا فإنه في المجتمعات غير المحدودة Infinit يستحيل إجراء حصر شامل لكل مفردات المجتمع في الوقت المتاح للدراسة ويكتفي في هذه الحالة بدراسة عينة بدون تكوين إطار للمعاينة. فمثلاً إذا كان بصدد اختيار عينة من أسر

أحد الأقسام الإدارية في مدينة ما وذلك لتقدير متوسط الدخل، فإنه يتحتم عليه اختيار عينة من إطار (أو قائمة) يحتوي على جميع أسر هذا القسم الإداري بالمدينة. ولا يجوز له في هذه الحالة أن يختار العينة من دليل التليفون مثلاً إذ أنه من المعروف أن مثل هذا الدليل لا يتضمن جميع أسر القسم الإداري قيد البحث.

الاختيار غير العشوائي (العمدي)

يلجأ الباحث أحياناً إلى اختيار مفردات عينة بطريقة غير عشوائية أو متعمدة، فمثلاً قد يختار أحد الباحثين عينة يرى أنها تمثل المجتمع بالنسبة إلى صفة أو خاصة ما دون غيرها وبعبارة أخرى يكون الأساس في الاختيار هو الباحث الذي يحدد بنفسه المفردات الداخلة في العينة متحيزاً لتفكيره ومعتمداً في تحديد المفردات. فمثلاً إذا أراد باحث دراسة مستوى المعيشة في الريف المصري فقد يعتقد أن قرية معينة في نظره تمثل مستوى المعيشة في كل الريف المصري، وفي هذه الحالة إذا اختار هذه القرية كأساس للدراسة فإن وجهة نظر الباحث تكون غير صائبة ولن يحالفها التوفيق في استخلاص النتائج، لأنه باختيارنا لهذه القرية يكون معتمداً أو متحيزاً في الاختيار لغرض الدراسة. ويمكن القول أن الباحث في تعمده في اختيار مفردات العينة إنما يتخلى من فكرة العشوائية في اختيار مفردات العينة عن مفردات المجتمع.

بمعنى أننا يجب ألا نقلل من أهمية الاختيار العمدي فربما يكون هو أفضل الطرق عن الاختيار في حالة إذا ما كان المطلوب اختيار عينة صغيرة لمجتمع كبير. فإذا كان المطلوب اختيار قرية واحدة لتمثيل القطر المصري كله فإنه يمكن اعتبار الاختيار العمدي هو أفضل الطرق رغم ما فيه من تحيز، وذلك لأن اختيار قرية واحدة بطريقة عشوائية قد يؤدي إلى خطأ كبير.

الاختيار العشوائي

على الرغم من سهولة اختيار أو سحب غير عشوائية من المجتمع كله إلا أن

ذلك له أضراره البالغة على دقة النتائج تبعاً لوجود تحيز من قبل الباحث في اختيار مفردات العينة لعدم توافر عنصر العشوائية في الاختيار ولضمان الحصول على المعاينة غير المتحيزة التي تعطينا تقديرات واستنتاجات يمكن تعميمها على جميع مفردات المجتمع الأصلي بأعلى دقة لتكاليف محددة، أو لا بد أن تختار العينة بحيث تكون ممثلة تمثيلاً صادقاً للمجتمع يكون ذلك على أساس نظرية الاحتمالات أي بواسطة سحب وحدات بالتتابع لكل منها احتمال معروف في الاختيار في السحبات المختلفة. ولكي تكون مفردات العينة ممثلة لكل مفردات المجتمع الأصلي بأقل أخطاء ممكنة فلا بد من استخدام أسلوب العشوائية غير المتحيز في الاختيار.

والأساس في أسلوب الاختيار العشوائي للعينة هو أن جميع مفردات المجتمع موضع الدراسة لها نفس الفرصة في الاختيار وهذا يعني عدم الاهتمام ببعض المفردات دون الأخرى وإتاحة الفرص المتكافئة أمام كل مفردة لتكون ضمن العينة. والمعنى الرياضي لكفاءة الفرص في الاختيار العشوائي هو أن يكون احتمال ظهور أي مفردة من مفردات المجتمع في العينة يساوي $\frac{\text{حجم العينة}}{\text{حجم المجتمع}}$ (أي $\frac{ن}{ق}$). وبذلك فإن الشرط الإحصائي الأساسي لاختيار مفردات العينة من بين مفردات المجتمع هو عدم التحيز في الاختيار حتى نضمن ـ إلى درجة ما ـ تمثيل العينة للمجتمع الذي نريد معاينته تمثيلاً صادقاً مع تقليل الأخطاء الأخرى المعاينة.

وتتم عملية اختيار مفردات العينة بالأسلوب العشوائي بإحدى الطرق الآتية:

أ ـ طريقة السحب العشوائي (القرعة)

عند اتباع هذه الطريقة تعطي مفردات المجتمع الأصلي أرقاماً مسلسلة تكتب على بطاقات متشابهة وبعد أن تخلط خلطاً جيداً. يكفي لإضاعة أي أثر للترتيب المتعمد ـ يسحب عشوائياً منها عدد من البطاقات يساوي عند مفردات العينة

المطلوبة. وتلائم هذه الطريقة سحب العينات من المجتمعات الصغيرة الحجم حيث لا تحتاج إلى مجهود كبير أو وقت طويل في عملية ترقيم مفردات المجتمع الأصلي وسحب العينة منه.

ب ـ طريقة الجداول العشوائية

يصعب لتباع طريقة السحب العشوائي في الاختيار في حالة المجتمعات الكبيرة الحجم حيث تحتاج عملية الترقيم إلى مجهود كبير كما أنها تكون مضيعة للوقت. ولذلك يفضل الرجوع إلى جداول خاصة تعرف باسم جداول الأرقام العشوائية Tables of Random Numbers وهي جداول أرقامها مختارة بطريقة عشوائية (أي أنها أرقام لا ترتبط ببعضها بأي أسلوب رياضي فهي لا تكون بينها متتالية عددية أو هندسية موضوعة في شكل أعمدة تتناسب مع حجم مجتمع إحصائي يتكون من أي عدد من المفردات كما يبدو من الجدول التالي (جدول رقم ١ ـ ١).

جدول رقم (١ ـ ١) جدول الأرقام العشوائية(١)

٢٠	١٧	٤٢	٢٨	٢٣	١٧	٥٩	٦٦	٣٨	٦١	٠٢	١٠	٨٦	١٠	٥١	٥٥	٩٢	٥٢
٧٤	٤٩	٠٤	٤٩	٠٣	٠٤	٠١	٣٣	٥٣	٧٠	١١	٥٤	٤٨	٦٣	٩٤	٦٠	٩٤	٤٩
٩٤	٧٠	٤٩	٣١	٣٨	٦٧	٢٣	٤٢	٢٩	٦٥	٤٠	٨٨	٧٨	٧١	٣٧	١٨	٤٨	٦٤
٢٢	١٥	٧٨	١٥	٦٩	٨٤	٣٢	٥٢	٣٢	٥٤	١٥	١٢	٥٤	٠٢	٠١	٣٧	٣٨	٣٧
٩٣	٢٩	١٢	١٨	٢٧	٣٠	٣٠	٥٥	٩١	٨٧	٥٠	٥٧	٥٨	٥١	٤٩	٣٦	١٢	٥٣
٤٥	٠٤	٧٧	٣٦	١٤	٩٩	٤٥	٥٣	٩٥	٦٩	٨٥	٠٣	٨٣	٥١	٨٧	٨٥	٥٦	٩٧
٤٤	٩١	٩٩	٤٩	٨٩	٣٩	٩٤	٦٠	٤٨	٤٩	٠٦	٧٧	٦٤	٧٢	٥٩	٢٦	٠٨	٥١
١٦	٢٣	٩١	٠٢	١٩	٩٦	٤٧	٥٩	٨٩	٦٥	٢٧	٨٤	٣٠	٩٢	٦٣	٣٧	٢٦	٢٤
٠٤	٥٠	٦٥	٠٤	٦٥	٦٥	٨٢	٤٢	٧٠	٥١	٥٥	٠٤	٦١	٤٧	٨٨	٨٣	٩٩	٣٤
٣٢	٧٠	٠١	٧٢	٠٣	٦١	٦٦	٢٦	٢٤	٧١	٢٢	٧٧	٨٨	٣٣	١٧	٧٨	٠٨	٩٢
٠٣	٦٤	٥٩	٠٧	٤٢	٩٥	٨١	٣٩	٠٦	٤١	٢٠	٨١	٩٢	٣٤	٥١	٩٠	٣٩	٠٨
٦٢	٤٩	٠٠	٩٠	٦٧	٨٦	٩٣	٤٨	٣١	٨٣	١٩	٠٧	٦٧	٦٨	٤٩	٠٣	٢٧	٤٧
٦١	٠٠	٩٥	٨٦	٩٨	٣٦	١٤	٠٣	٤٨	٨٨	٥١	٠٧	٣٣	٤٠	٠٦	٨٦	٣٣	٧٦
٨٩	٠٣	٩٠	٤٩	٢٨	٧٤	٢١	٠٤	٠٩	٨٦	٦٠	٤٥	٢٢	٠٣	٥٢	٨٠	٠١	٧٩
٠١	٧٢	٣٣	٨٥	٥٢	٤٠	٦٠	٠٧	٠٦	٧١	٨٩	٢٧	١٤	٢٩	٥٥	٢٤	٨٥	٧٩
٣٧	٥٦	٤٩	٧٩	٣٤	٣٤	٣٢	٢٢	٦٠	٥٣	٩١	١٧	٢٣	٢٦	٤٤	٧٠	٩٣	١٤
٤٩	٠٥	٧٤	٤٨	١٠	٥٥	٣٥	٢٥	٢٤	٢٨	٢٠	٢٢	٣٥	٦٦	٦٦	٣٤	٢٦	٣٥
٤٩	٧٤	٣٧	٢٥	٩٧	٢٦	٣٣	٩٤	٤٢	٢٣	٠١	٢٨	٥٩	٥٨	٩٢	٦٩	٠٣	٦٦
٢٠	٢٦	٢٢	٤٣	٨٨	٠٨	١٩	٨٥	٠٨	١٢	٩٧	٦٥	٦٥	٦٣	٥٦	٠٧	٩٧	٨٥
٤٨	٨٧	٧٧	٩٦	٤٣	٣٩	٧٦	٩٣	٠٨	٧٩	٢٢	١٨	٥٤	٥٥	٩٣	٧٥	٩٧	٢٦
٠٨	٧٢	٨٧	٤٦	٧٥	٧٣	٠٠	١١	٢٧	٧	٠٥	٢٠	٣٠	٨٥	٢٢	٢١	٠٤	٦٧
٩٥	٩٧	٩٨	٦٢	١٧	٢٧	٣١	٤٢	٦٤	٧١	٤٦	٢٢	٣٢	٧٥	١٩	٣٢	٢٠	٩٩
٣٧	٩٩	٥٧	٣١	٧٠	٢٠	٤٦	٥٥	٤٦	١٢	٢٤	٣٢	٣٦	٧٤	٦٩	٢٠	٧٢	١٠
٠٥	٧٩	٥٨	٣٧	٨٥	٣٣	٧٥	١٨	٨٨	٧١	٢٣	٤٤	٥٤	٢٨	٠٠	٤٨	٩٦	٢٣
٥٥	٨٥	٦٣	٤٢	٠٠	٧٩	٩١	٢٢	٢٩	٠١	٤١	٣٩	٥١	٤٠	٣٦	٦٥	٣٦	١١

(١) يحسن الرجوع إلى الجداول المفصلة في:

Lindley and Miller (1953). Cambridge Elementary Statistical Tables. Cambridge.

١٩	٠٨	٨٦	٦٤	٦٩	٠٥	٢٤	٤٩	٨٥	٠٣	٧٢	٢٤	٣٦	٦٨	٢٥	٩٦	٢٨	٦٧
٨٧	٠٦	٥٧	٨٩	٦٠	٤٥	٨٤	٧٣	٤٣	٨٦	٨٢	٥١	٥٩	٣٢	٧٨	٩٤	٨٦	٨٥
	٦٢	٧٥	٤٧	١٨	١١	٣٧	٢٥ ٥٨	١٣	٦٣	٤٤	٧٨	٨٨	٥	٩	٦٠	١٠	١٠
١٩	٢٠	٢٠	٦٦	٤٧	٢٣	٤٤	٨٧	٨٩	٢٦	٨٠	٧٧	٨٠	٩٠	٤٨	٨٩	٥٥٤	٩٤
٠٩	١٤	٣٧	٥٦	١٥	٩٤	٣	٢٩	١٩	٦٢	٦٢	٣٣	٢٠	٢٣	٧٧	٧٧	٦٣	١١
٦١	٠٣	٢٥	٢٤	٠٤	٢٦	٤٠	٦٨	٦٢	٩٤	٥٧	٣٨	٥٥	٥٤	٠٤	١٦	٠٠	٦٤
٩٦	٧٠	٩٨	٤٥	٢١	٣٠	٤٦	٨٨	٦٠	١٠	٥٨	٦٠	٤١	٧٨	٢٣	١٣	٩٤	٥٠
٨٥	١٧	١٩	٢٧	٩٧	٤٨	٨٥	٧٥	٥٩	٣٤	٠٥	٤٥	١٣	٤٤	٩٦	٣٧	٩٨	٦٦
٨٠	٠٨	٢٨	٧٢	٦٦	٥٧	٦٢	٧٩	٩٠	٦٠	٩١	٩٠	٧٧	٦٠	٨٣	٤٢	٩١	٦٦
٤٣	٨٥	٨٤	٥٨	٤٢	٩٠	٤٥	٣٩	٢٩	٢١	٤٠	١٩	٠٧	٠٢	١٨	١٢	٥٨	٢٣
٢٥	٢٧	٣٠	٠٥	٢٤	٩٨	٠٤	٠٢	٩٠	٥٠	٠٥	٧٣	٤٠	٧٢	٠٦	٤٠	٤٩	٥٢
٤١	٨٤	٧٦	٨٦	٢٣	٤٠	٥٨	٤٥	٦٦	٦٦	٤٧	٧٩	٣٠	٧٨	٩٩	٩٣	٩٨	٧٤
٤٤	٩٣	٦٩	٠٥	٢١	٢٣	٨٨	٦٩	٤٥	٥٤	٥٢	٦٩	٨٨	٠١	٣٠	٥٤	٢٦	٥٠
٣٤	٥٦	٥٦	٣٩	٧٣	٢٨	٣٥	١٩	٣٤	٢٨	٨٤	٩٦	٧٩	٣٣	٨٩	٦١	٤٦	٤٩
٩٠	٦٧	٧٦	٨٦	٦٦	٢٥	٠٠	٣٦	٠٣	٢٢	٨٨	٧٣	٣٩	٧٨	٤٤	١٣	٦٥	١٩
٣٨	٢٣	٨٦	٥٥	٢٨	٢٨	٠٠	١٤	٦١	٧٢	٤٠	٨١	٥٩	٨٧	٦٧	٤٧	١٧	٦٤
٦٠	٤٢	١٧	٤٨	٨٩	١١	٨٨	٠١	٩٥	٥٧	٥٦	٥٦	٩٧	٦٨	٣٧	٩٧	٤٣	١٨
٠١	٣١	٧٥	٤٤	٨٨	٦٦	٨١	٦٦	٥٣	١٥	٦١	٩٦	٠٩	٥١	٨٧	٦٠	٥٨	٦٥
٥٣	٣٤	٧٨	٦٤	٩٣	٤٥	١٥	٤٣	٧٥	٣١	٦٥	٨٥	١٤	٩١	٠٠	٣١	٩٠	٧٩
١١	٣٤	٦٧	٦٥	٧٦	٦٣	٤٠	٢٠	٤٢	٩٤	٠٩	١٦	٠٥	٥٩	١٥	٠٠	٢٣	٠٧
٥١	٤٠	٢٨	١٩	١٤	٠٠	١٩	٣٣	٣٢	٣٠	٤٦	٩٥	٥١	٠١	٢٤	١٤	٠٨	٩٠
٣٠	٥٦	٩٧	٠٠	٣٠	٦٥	٨٥	١٢	٠٣	٤١	١٣	٣٤	٨٢	٢١	٠٢	٦٢	٨٢	٥٣
٦٤	٥٩	٣٣	٢٣	٣١	٣٩	٨٤	٥٤	٣٣	٢٠	٢٥	٧٦	٥٠	٠٤	١٥	٢٦	١٧	٩٨
٤٤	٥٩	٣٥	٨٩	١٧	٦٥	٥٤	٦٤	٥٠	٤٥	٦٢	٣٠	٤٠	٨٢	٤٤	١٢	٩١	٠٨
٥٠	٩٠	٧٤	١١	٤١	٣٣	٣٧	٩٦	٥٤	٨٥	١٦	١٧	٨٧	٨٤	٧٧	٤٦	٢١	٣٧

وعند استخدام هذه الجداول تعطي أرقام مسلسلة لمفردات المجتمع الذي نريد معاينته ثم يختار اختياراً عشوائياً بداية تؤخذ من عندها الأرقام التي تدخل ضمن مفردات العينة مع استبعاد الأرقام المتكررة أو الأرقام التي تزيد عن حجم المجتمع. ولتحقيق ذلك يلزم أخذ أعمدة تحتوي على عدد من الخانات تساوي عدد أرقام حجم المجتمع. ولشرح ذلك نقول لو كانت بيانات عن مجتمع يتكون من ١٠٠ مفردة فبإمكاننا اختيار الأرقام بالطريقة العشوائية من العمودين الأول والثاني من الجدول رقم (١ ـ ١). فإذا كنا نريد اختيار عينة مكونة من ١٠ مفردات من ١٠٠ مفردة، فإننا سنرى أن أرقام المفردات المطلوبة من الجدول السابق، هي ٢٠، ٧٤، ٩٤، ٢٢، ٩٣، ٤٥، ١٦، ٤، ٣٢. أما إذا كان حجم المجتمع كبيراً كان يتكون من ١٠٠٠٠ مفردة، ففي هذه الحالة تستعمل الأعمدة الأربعة في الجدول المذكور. فإذا أردنا سحب عينة مكونة من ١٠٠ مفردة من هذا المجتمع، فإن أرقام مفردات العينة المختارة من الجدول بهذه الطريقة تكون هي على التوالي ١٧٢٠ ثم تليها ٤٩٧٤ وبعدها ١٠٩٤ وهكذا حتى يصل عدد المفردات ١٠٠ وهي عدد مفردات العينة المطلوبة. وإذا كنا بصدد معاينة ٢٠٠٠ مفردة فقط فإننا نختار من الجدول السابق أرقام مكونة من أربعة حدود لمفردات العينة المطلوبة وهي ١٠٠ مفردة وتتم عملية الاختيار بإحدى الطريقتين: أما أن تختار الأرقام العشوائية من العمودين الأول والثاني كمفردات العينة بقبول أي رقم يقع بين أو ٢٠٠٠ ورفض أي رقم آخر يقع بين ٢٠٠٠ و ٩٩٩٩، ونستمر في عملية القبول والرفض حتى يصل عدد المفردات ١٠٠ مفردة. ويعاب على هذه العملية أنها تستنفذ وقتاً طويلاً وذلك بسبب ارتفاع معدل رفض الأرقام التي تزيد عن الحجم الكلي للمفردات الذي نريد سحب العينة منه. ولهذا السبب تستبدل طريقة قبول ورفض الأرقام العشوائية بطريقة أخرى يمكن بها اختيار المفردات المطلوبة وذلك على أساس قبول جميع الأرقام التي تزيد عن رقم ٢٠٠٠ واعتبارها تكرارات الأرقام بين ١و٢٠٠٠، بمعنى أن الأرقام العشوائية بين ٢٠٠١ و ٤٠٠٠، ٤٠٠١ و ٦٠٠٠، ٦٠٠١و٨٠٠٠، ٨٠٠١ و ١٠٠٠٠ يمكن اعتبارها أرقاماً إضافية (جديدة) للأرقام من ١ إلى ٢٠٠٠ وتمتاز هذه الطريقة بأن فيها اختصار كبير للوقت في اختيار أرقام

المطلوبة من جميع الأرقام المحصورة بين ١ و ١٠٠٠٠

وفي بعض الحالات تكون البيانات المتاحة عن مجتمع الظاهرة قيد البحث في شكل مجموعات ويراد أخذ عينة عشوائية من المجموعات ككل وليس من كل مجموعة على حدة فمثلاً إذا كانت لدينا بيانات عن أعداد السكان لعدد كبير من المراكز العمرانية التي تتكون من عدة مجموعات قرى صغيرة، قرى كبيرة، مدن صغيرة، مدن كبيرة (عواصم المحافظات) فإنه بإمكاننا أخذ عينة واحدة من كل هذه المجموعات باستخدام الطريقة العشوائية في الاختبار. ويتم ذلك بأن ترتب أعداد كل المجموعات في قائمة حتى نحصل على المجموع الكلي لأعداد جميع المراكز العمرانية المطلوب معاينتها. فمثلاً إذا كان عدد المراكز العمرانية للمجموعات الأربع هو: ١٣٦، ٦٧، ٣٢، ٧ على الترتيب، فإنها تأخذ ترقيماً من ١ إلى ١٣٦، ومن ١٣٧ إلى ٢٠٤، ومن ٢٠٥ إلى ٢٣٧، ثم من ٢٣٨ إلى ٢٤٥ على الترتيب أيضاً. وبالتالي يتكون المجتمع الإحصائي للمجموعات الأربع من ٢٤٥ مركزاً عمرانياً والذي يمكن منه سحب عينة بالطريقة العشوائية السابق شرحها.

جـ ـ طريقة السحب بواسطة الحاسب الآلي

تستخدم هذه الطريقة في حالة سحب عينات كبيرة الحجم من مجتمعات تتميز بأحجام كبيرة جداً. وتحقق هذه الطريقة درجة عالية من العشوائية وعدم التحيز حيث أن الباحث لا يتدخل في عملية الاختيار.

ويطبق هذا الأسلوب حالياً في الأبحاث العلمية التي تجري في معظم دول غرب أوروبا والولايات المتحدة الأمريكية، وفي سبيله للانتشار في الدول الأخرى التي أخذت على عاتقها حديثاً، تطوير أجهزتها الإحصائية بإدخال الحاسبات الآلية Computers ومن بينها جمهورية مصر العربية

وبالإضافة إلى السحب الآلي لمفردات العينة، فإنه من الممكن الآن تغذية الحاسب الآلي ببرامج يمكن بواسطتها أن يضع جداول للأرقام العشوائية التي من أمثلتها الجدول الحالي

جدول رقم (١ ـ ٢) جدول الأرقام العشوائية
بواسطة الحاسب الآلي(١)

٧٣٢٤٤٤٧٠١١٤٠٢٧٠٤٤٠١٢٤٤٥٥٩٩٠٣٣
١١٣٢٤٠٣٩٥٢٤٠٧٦٢١٣٢٠٩١٤١٣٣٧٥٧٨٠
٧٦٢٤٥٢٢٣٢١٣٤٧٣٧٦٨٩٨٨٥٨٤٤٦٢٣٨٣٤
٣٠٩٠٧٢٤٣١٩٢٣٧٣٣٧١٩٥٣١٠١٤٢٨٤٤٢٨٤
٢٩٢٦٢٤٣٤٧٢١٣٤٨٥٦١٤٢٠١٢٢٨٤٦٣٠٧٢
٠٤١٤١٥٩٧٧٨٦٤٣٣٨٠٨٩٤٤٤٧٢١٢١٣٣٢
٤٣٨٦١٦٨٤٥٣٢٤٣٠٣٠٤٦١٨٣٧٨٣٤٣١٨٥٦
٢٠٣٧٠١٢٧١٣٩٨٠٨٥٨٧٩٧١٧٤٠٣٠٨٤٩٦١
٧١٤٩٦٨٥٥٥٩٢٩٧٥٦٨٢٢٧٢٤٦٣٤٦٦٧٧٨٧
٩١٧٢٤٦٩٩٤٧٣١٠٥٧٤٦٣٣٨٦٧٢٠٢٩١٤٥٣٩
٣٨٠٤٣٢٠٣٥٩٧٧٧٤٢٦٦١٥٩٩٥١٢١٤١٩٥٤
٦٩٥٦٥٦٤٠٣٤١٠٩٢٠٩٩٥١٦١٥٨٦٠٠٣٨٨١٨
٣٧٨٥٣٤٦٤٠٢٣٨٨٤٨٩٢٩٠٣٥٤٧٩٢١٥٣٧٠
٠٥٦٧٧٠٩٥٢٩٤٢١٧٤٠٢٤٣٣٢٥٣٣٧٥٠١٢١
٣٧٩٦٥٩٤٣٤٩٦٠٧٥٨٣٠٧٥٧٦٨٩٠٧٨٣٣٩٢
٩١٢٤٥٢١٠٩٤٨٥٩٠٤٩٧٩٩٦٠٧٠٦٤٨٩٩٤٨

(١) وضع برنامج الحاسب الآلي للحصول على أرقام هذا الجدول Dr M.Mc Cullagh أستاذ الكارتوجرافيا والتحليل الكمي في الجغرافيا ـ بقسم الجغرافيا ـ جامعة نوتنجهام ـ إنجلترا

٧٦٧٧٢٤٤٢٠٢ ٦١ ٥٩١٦٨٤٢ ٤٩ ٢٥٨٢
٢١٨٩٧٦٨٨٨٧٩٥٩٩٩٦٦ ٥٥١٧٥٦٧٩٦٦١٢٧
٧٦٢٩٩٤٣٤٢٩٢٢٥٣٦٧٢٠٥٩٥٥٥٦١٢٩٥٣
١٧٥٦٧٩٦٦٦٢٧٧٦٧٧٢٤٤٩١١ ٦٨٧٦١٧٨٧٨
٥٥٥٦١٣٩٥٣٠٢١٨٩٧٦٨٨٨٧٩٥٩٩٩٦٦٠٥٥
٠٦٨٧٦١٧٨٧٨٧٦٢٩٩٤٣٤٢٩٢٢٥٣٦٧٢٠٥٩
٦٦٠٨٠٥٨٠٥٥١٧٥٦٧٩٦٦٢٧٧٦٧٧٢٤٤٩١١
١٤٥٤٣٤٩١١٤٨٦٨٥٨٩٣٧٧٦٠٢٠٦١٩٨٢٥٤
١٨٢٤٢١٤٩٦٤٨٤٩٨١٤٥٢٧٧٥٦٦٧٧٥٠١٢٦
٣٩٣٤٠٨١٩٤٥٩٤٥٥٧٥٩٨٠٢٥٧١٤٦٥٨٠٦٠
٢٢١٣٨٩٥١٢٦٧٧٤٧٩٩١٣٧٢٥٠٨٣٤٩٩٦٠٣
٢٨٣٠٣٦٢٦٦٢٦٦٢٠٣٠٧١٢١٢٢٩٣٩١٤٤٦٠٢
٤٦٥٦٩٥١٩٩٢٢٤٦٣٦٨٣٢٣٤٧٨١٣٣٦٨٧٢٩
٥٩٩٠١٧٥٣٩٢٣٠٢١٣٤٦٣٢١٩٤٥٦١٧٥٠٩٦
٨٢٠٢٠٢٧٥١٩٧٥٣٦٨٥٢٧٢١٠٨٤٢٢١٣٨٣٥
٣٨٣٧٨١١٩٧٦٩٦٣١٤٢٣٤٣٩٠٠٤٠٧٤٤٣١٤
٠٥٠٦٦٥٥٦٣٨٠٨٤١٦٥٩٧١٥٢٣٥٢٩١٩٧١٨
٤٠٦٢٧٥٢١٠٣٣٨٣٢٦٤٢٠٣٢٣٧٤٥٠١٠٦٨١
٣٦٩٠٨١٣٦٥٥٩٠١٢٣٦٧٩٠٣٥١٦٩٧٢٣١٥٨

نلخص من كل ما سبق أنه على الرغم مس أن الباحث يجب أن يكون حذرا وعير منحيراً عند احتياره لمعردات العينة بإحدى طرق الاختيار السابقة، إلا أن هناك أنواع كثيرة مس الأخطاء التي تصاحب أسلوب سحب العينات يكون مصدرها

الرئيسي أما تحيز الباحث (خطأ التحيز Bias Error) أو تحيز البحوث، أو عدم التزام الباحث بأسلوب العمل الإحصائي أي عدم اتباع القواعد السليمة في جمع المعلومات أو سوء التقدير والإهمال في العمل وإلى جانب ذلك هناك أيضاً خطأ الصدفة أو الخطأ العشوائي Random Error الذي يعد من أهم أخطاء أسلوب المعاينة في الدراسة وكما سبق أن ذكرنا أن خطأ الصدفة من الأخطاء التي تخرج عن نطاق القصد أو التعمد حيث أن قيمة الخطأ تتفاوت من عينة لأخرى، وأن مصدره الأساسي هو الاعتماد على بيانات العينة فقط في استخلاص النتائج الخاصة بالمجتمع الذي تمثله هذه العينة (راجع الفصل الأول).

مثال تطبيقي لكيفية سحب عينة

لنفرض أنه لدراسة تباين الإنتاج الزراعي في إحدى المحافظات وتأثيرها على الدخل الزراعي في المجتمع الريفي عن طريق اختيار الفرق بين متوسطات إنتاج الأرض الزراعية في قرى المحافظة والتي يبلغ عددها ٩٥ قرية، ولما كان من الصعب دراسة القرى كلها لكثرة عددها أو لصعوبة الوصول إلى بعض منها، فقد تقرر أخذ عينة من ١٠ قرى. والمطلوب تحديد هذه القرى العشرة من بين القرى كلها. ولتحقيق ذلك تستخدم طريقة الجداول العشوائية لاستخراج أرقام القرى المطلوبة باتباع الخطوات التالية: ـ

١ ـ إعطاء كل قرية رقماً خاصاً بها من ١ إلى ٩٥.

٢ ـ بالرجوع إلى جداول الأعداد العشوائية (جدول رقم ١ ـ ١) يمكن اتخاذ أي عمود أو صف واختيار عشرة أرقام منه.

٣ ـ إذا أخذنا الصف اورل من الجدول فإن الأرقام العشوائية للقرى المختارة تكون هي: ٢٠، ١٧، ٤٢، ٢٨، ٢٣، ٥٩، ٦٦، ٣٨، ٦١، ٢ (لاحظ أننا لم نأخذ الرقم السادس في الصف وهو رقم ١٧ لأنه مكرر واستبدلناه بالرقم ٢ من نفس الصف).

٤ ـ تكون الأرقام العشرة السابق هي الأرقام العشوائية الممثلة لمجتمع القرى

(٩٥) أما إذا كان عدد القرى هو ٩٥٠ قرية وأريد اختيار عينة بحجم ما فإن أرقام مفردات العينة ستحتوي في هده الحالة على أرقام مكونة من ثلاثة حدود فإذا كانت العينة مكونة من ٣٠ قرية مثلاً فإننا نأخذ العمودين الأول والثاني من جدول الأرقام العشوائية ونختار منها الأرقام المكونة من ثلاثة حدود، والتي تمثل مفردات العينة فتكون هي. -

٧٢٠، ٠٩٤، ٥٢٢، ٤٤٥، ١٤٤، ٣١٦، ٠٠٤، ٠٣٢، ٤٠٣، ٠٦١، ٣٨٩، ٢٠١، ٦٢٧، ٥٤٩، ٤٤٩، ٦٢٠، ٧٤٨، ٢٠٨، ٧٩٥، ٩٣٧، ٩٠٥، ٥٥٥، ٨٦٧، ٦٨٥، ٠٤٠، ٥٨٤، ٣١١، ٠٦٤، ٤٥٠، ٨٨٦.

ويلاحظ أننا رفضنا الأرقام ٩٧٤، ٩٩٣، ٩٦٢ نظراً لأنها أكبر من حجم المجمتع الأصلي (٩٥٠ قرية) واخترنا بدلاً منها الأرقام العشوائية الثلاثة الأخيرة ٠٦٤، ٤٥٠، ٨٨٦.

(٣) تحديد نوع العينة

يجمع كثير من الإحصائيين والباحثين على أن تحديد نوع العينة المختارة التي يجب أن تتوفر فيها صفة إعطاء نتائج ذات دقة معينة بأقل تكاليف ممكنة أو بأعلى دقة بتكاليف محددة يتوقف على طبيعة الدراسة ونوعية وتركيب المجتمع الذي ستسحب منه العينة والوسيلة أو الأداة المستخدمة في جمع البيانات، ووجهة نظر الباحث نفسه.

ويمكن تصنيف العينات على أساس عامل العشوائية في الاختيار إلى قسمين رئيسين: القسم الأول يشمل العينات العشوائية التي يعتمد الباحث في تصميمها على نظرية الاحتمالات في إعطاء الفرص المتكافئة لمفردات المجتمع لأن تظهر في العينة، أما القسم الثاني فيتضمن العينات العمدية (غير العشوائية) والتي يكون فيها تحيز الباحث واضحاً في اختيار مفردات العينة وذلك عن طريق إعطاء فرص غير متكافئة للمفردات نتيجة تعمده اختيار بعض المفردات دون غيرها من مفردات

المجتمع الذي يريد معاينته ولكل من القسمين أنواع متعددة ومتنوعة من العينات سندرسها بالتفصيل كما يلي: -

أولاً: العينات العشوائية Random Samples

يستخدم أحياناً مصطلح العينات الاحتمالية للدلالة على العينات العشوائية وذلك لأنها تعتمد كما سبق القول على نظرية الاحتمالات في اختيار مفردات العينة من بين مفردات المجتمع عن طريق سحب تلك المفردات بالتتابع لكل منها احتمال معروف في الاختيار في السحبات المختلفة. وبعبارة أخرى تجري المعاينة العشوائية (الاحتمالية) حسب خطة إحصائية لا يسمح فيها للباحث، أو حتى المفردات في العينة، التدخل في اختيار العناصر الخاصة بالبيانات المراد جمعها أو أن يستعاض عن بعض المفردات التي يجب اختيارها ومنها للعينة بمفردات أخرى أسهل في حالة صعوبة الوصول إلى، أو الحصول على بيانات المفردات الأولى. وسوف نعرض فيما يلي لأهم أنواع العينات العشوائية وهي: العينة العشوائية البسيطة، العينة العشوائية المنتظمة، العينة العشوائية الطبقية، والعينة العشوائية المتعددة المراحل.

أ ـ العينة العشوائية البسيطة Simple Random Sample

يلائم هذا النوع من العينات الدراسات التي تهدف إلى تحديد خصائص المجتمع الذي تمثل مفرداته مجموعة متجانسة أي ذات نوعية واحدة، مثل مجموعة من الطلاب عند مستوى عمري متقارب وأوزان متساوية تقريباً، أو مثل إنتاج أحد المصانع المعلبات لوحدات إنتاجية (عبوات) ذات أوزان واحدة. ويتم اختيار مفردات العينة على أساس تساوي فرص الاختيار المستقل أمام كل مفردات المجتمع،

ب ـ العينة العشوائية المنتظمة Systematic Random Sample

يستخدم هذا النوع من العينات عند دراسة المجتمعات التي تكون مفرداتها متخذة شكل انتظام متسق (أي تتصف بعدم التغير أو التباين الشديد ومتدرج من حيث التنوع). وفي هذه العينة تأكيد على تسلسل المفردات وفقاً للتنوع في الخصائص المميزة للمجتمع الأصلي ويتبع في اختيار مفردات العينة الأسلوب العشوائي، كما في العينة العشوائية البسيطة غير أن الاختيار يتم بطريقة منتظمة، أمام جميع المفردات للظهور في العينة هي أساس الاختيار، وعندئذ تنتهي العشوائية ويبدأ اختيار مفردات العينة وفق نظام أو قاعدة معينة حتى نحصل على العينة المطلوبة.

وتمتاز هذه العينة بانتظام وثبات الفترات أو التباعد بين وحدات أو مفردات العينة. وفيها نبدأ بتحديد مقدار تمثيل كل مفردة من مفردات العينة لعدد معين من مفردات المجتمع، ثم نقوم باختيار المفردة الأولى (أو أحد أعداد مقدار أو نسبة التمثيل) عشوائياً ونضيف إليها مقدار التمثيل بطريقة منتظمة حتى نحصل على بقية مفردات العينة بشكل منتظم وبتباعد متساوي فيما بينها. ولتحديد مقدار التمثيل نقسم عدد مفردات المجتمع الأصلي على حجم العينة المطلوب. فمثلاً إذا كان حجم المجتمع الأصلي ٦٠٠٠ مفردة، وأردنا اختيار عينة من ٢٠٠ مفردة، فهذا يعني أننا نريد اختيار مفردة واحدة لكل ٣٠ مفردة (أي $\frac{٦٠٠٠}{٢٠٠}$ = ٣٠). وأحد الطرق المتبعة هي، أن نختار عدداً عشوائياً بين ١، ٣٠ ولو فرضنا أن هذا الرقم هو ٢٢، فإنه يمكن تحديد مفردات العينة مباشرة بإضافة مقدار التمثيل (٣٠) بطريقة منتظمة الى الرقم (٣٢) وذلك على النحو التالي: ـ

٢٢، ٥٢، ٨٢، ١١٢، ١٤٢، ١٧٢، ٢٠٢، ٢٣٢، ٢٦٢، ٢٩٢... وهكذا حتى نصل إلى المفردة الأخيرة ويكون رقمها ٥٩٩٢ وهناك طريقة لإيجاد ترتيب أي مفردة من مفردات العينة وذلك باستخدام المعادلة الآتية ـ

رقم المفردة = الرقم العشوائي المختار + (ترتيب المفردة ـ ١) × مقدار التمثيل

أي أن:

ب = (ش + (ت ـ ١) × ث) .... (٢ ـ ١٣)

فإذ أردنا مثلاً معرفة ترتيب المفردة رقم ١٠٠ من ٢٠٠ مفردة في العينة السابقة فإن:

ب = ٢٢ + (١٠٠ ـ ١) × ٣٠ = ٢٩٩٢

وتتميز العينة المنتظمة بأنها أسهل وأسرع في التطبيق من العينة العشوائية البسيطة، إذ أنها لا تحتاج إلى اختيار كل المفردات بالطريقة العشوائية والذي قد ينتج عنه تكرار سحب بعض المفردات، كما أنها تمثل توزيعاً متسقاً (منتظماً) Uniform للمجتمع الذي تسحب منه العينة بعكس العينة العشوائية البسيطة التي ينتج عنها في معظم الأحيان توزيعات مكانية تتخذ أشكالاً عنقودية Clusters or Bunching أو تتباعد فيها المفردات بعضها عن بعض تباعداً لا يمكن تقليله إلا بزيادة حجم العينة.

ويعاب على العينات المنتظمة إذا أجريت على التوزيعات المكانية أنهاتؤكد صفة الانتظام والاتساق في الظاهرة أو الظاهرات الجغرافية موضع المعاينة والتي تكون غالباً إن لم يكن دائماً، في حالة تغير وتطور مستمرين. كما قد تتصف المعاينة بأنها لا تعطى عينة غير متحيزة وممثلة للمجتمع الذي سحبت منه، بسبب خطأ التحيز وعدم اتباع أسلوب تكافؤ الفرص في اختيار مفردات العينة، إلا إذا كانت مفردات المجتمع الأصلي موزعة توزيعاً عشوائياً ـ وكثيراً من المجتمعات الإحصائية للظاهرات الجغرافية لا تتصف بصفة العشوائية في توزيعاتها. وعلى العموم فإن خطأ التحيز إن وجد في بيانات العينات المنتظمة فإن قيمته تكون صغيرة جداً بحيث يمكن إهمالها عند تطبيق أساليب التحليل الإحصائي الكمي على هذه البيانات.

جـ ـ العينة العشوائية الطبقية Stratified Random Sample

يستخدم هذا النوع من العينات العشوائية في دراسة المجتمعات التي تتميز بتباين نوعيات مفرداتها والتي يمكن تقسيمها إلى مجموعات (أو طبقات Strata) لكل مجموعة أو طبقة منها خصائص ومميزات معينة، ولكنها تتباين فيما بينها تبايناً عظيماً في هذه الخصائص والمميزات. كما يتسم هذا النوع من العينات بأنه أكثر دقة في الاختيار العشوائي من العينات العشوائية البسيطة حيث أن المفردات الكلية للعينة الطبقية تكون أكثر تمثيلاً لجميع مجموعات أو طبقات المجتمع الأصلي مما يؤدي إلى تقليل في الأخطاء عند تعميم نتائج العينة على كل المجتمع. وهكذا تقوم العينة الطبقية في أساس تقيم المجتمع الأصلي إلى طبقات ثم تأخذ عينة عشوائية من كل طبقة بشكل يتناسب مع مفردات أو حجم تلك الطبقة. ولهذا فإن معاينة أية طبقة تتطلب عدة إجراءات تختلف عن تلك التي تتطلبها الطبقة أو الطبقات الأخرى. فمثلاً في التعدادات بالعينة التي تجري لحصر أعداد السكان في منازل أحد الشوارع بمدينة ما تختلف عن مثيلتها لحصر عدد العاملين في مؤسسة أو شركة ما. وبصفة عامة يمكن تلخيص طريقة اختيار العينة الطبقية في الخطوات الآتية: ـ

١ ـ تقسم المجتمع إلى مجموعات مميزة أو فئات فرعية (مجتمعات صغيرلا) متجانسة تعرف بالطبقات Strata.

٢ ـ تحديد نسبة مفردات كل مجموعة أو طبقة Stratum بالنسبة للعدد الكلي لمفردات المجتمع الأصلي (أي $\frac{\text{حجم الطبقة}}{\text{حجم المجتمع الأصلي}}$).

٣ ـ تحديد عدد مفردات العينة المطلوبة من كل طبقة، أو ما يعرف بالعينة الفرعية Sub sample التي تتحدد عن طريق نسبة حجم كل طبقة في المجتمع الأصلي والحجم الكلي للعينة

٤ ـ استخدام الأسلوب العشوائي لاختيار المفردات من كل طبقة

مثـال

أراد باحث سحب عينة حجمها ٥٠٠ عاملاً من مجتمع عمالي لأحد الصناعات تبلغ حجمه ٥٠٠٠ عاملاً وذلك حسب الحالة التعليمية، فقام بتقسيم عمال المجتمع إلى فئة من العمال الأميين وعددم ١٠٠٠ عامل وفئة من العمال الحاصلين على شهادات أقل من المتوسطة، ١٥٠٠ عاملاً، وفئة من العمال الحاصلين على شهادات متوسطة وفنية وعددهم ٢٥٠٠ عاملاً فكم من العمال من كل فئة يمكن اختيارها حتى يحصل على الحجم الكلي للعينة؟ ويتم ذلك بتحديد حجم العينة الفرعية عدد المفردات المطلوبة الممثلة لمجتمع الطبقة وهو يساوي:

$$\text{حجم العينة} \times \frac{\text{حجم الطبقة}}{\text{حجم المجتمع الأصلي}}$$

وبالتالي فإن:

عدد المفردات من الفئة (عمال أميين) $= ٥٠٠ \times \frac{١٠٠٠}{٥٠٠٠}$

$= ١٠٠$ عامل

عدد المفردات من الفئة (عمال الشهادات أقل من المتوسط) =

$٥٠٠ \times \frac{١٥٠٠}{٥٠٠٠} = ١٥٠$ عامل

عدد المفردات من الفئة (عمال الشهادات المتوسطة والفنية) =

$٥٠٠ \times \frac{٢٥٠٠}{٥٠٠٠} = ٢٥٠$ عاملاً

ويكـون عدد العمال المطلـوب اختيارهـم من كل فئـة من الفئـات الثلاث هو على الترتيـب ١٠٠، ١٥٠، ٢٥٠ ومجموعهـم لا بد وأن يسـاوي الحجم الكلـي للعينة المطلوبـة ٥٠٠ عـامـلاً من المجتمـع الأصلي لعمـال تلـك الصناعة.

وتجدر الإشارة هنا إلى أن الطبقية يمكن أن تكون طبقية طولية أو عرضية (في حالة مجتمع المساحات)، أو طبقية أفقية أو رأسية (في حالة نوع المناطق السكنية أو مستويات الدخل أو فئات العمر . . .).

أمثلـة تطبيقيـة

مثـال (١)

لدراسة أحد مظاهر النشاط الصناعي مثل صناعة الأحذية في منطقة تتميز بتباين توزيع منشآت أو ورش الصناعة في مختلف المراكز العمرانية لهذه المنطقة فإذاً أريد معاينة توزيع هذه الصناعة بطريقة المعاينة الطبقية فإننا نقوم بتقسيم المراكز العمرانية في المنطقة حسب حجم المكان بها إلى أربع مجموعات: قرى صغيرة، قرى كبيرة، مدن صغيرة، مدن كبيرة، ونختار عينة عشوائية من المراكز العمرانية لكل مجموعة نسبة معاينة $\frac{١}{١٠}$ (أي $\frac{\text{حجم العينة}}{\text{حجم المجتمع}} = \frac{\text{ن}}{\text{ى}} = \frac{١}{١٠}$) والتي تتناسب وحجم كل مجموعة من المجموعات الأربع، كما يتناسب المجموعة الكلي للعينات مع الحجم الكلي لمجتمع المراكز العمرانية). وتوضح ذلك البيانات التالية:

المجموعة	أ ـ عدد الوحدات (المراكز العمرانية في العينة	ب ـ العدد الكلي للوحدات في كل مجموعة	جـ ـ عدد ورش صناعة الأحذية في كل وحدات العينة	د ـ متوسط ورش العينة (كل وحدة) جـ / أ	هـ ـ تقدير تقدير المجموع الكلي لعدد الورش في كل مجموعة (ب×د)
قرى صغيرة	١٢	١٢٠	٢٤	٢	٢٤٠
قرى كبيرة	١٤	١٤٠	٣٥	٥ر٢	٣٥٠
مدن صغيرة	٥	٥٠	٦٠	١٢	٦٠٠
مدن كبيرة	٢	٢٠	٨٠	٤٠	٨٠٠
المجموع الكلي	٣٣	٣٣٠	١٩٩	٠٣ر٦	١٩٩٠

من البيانات السابقة الخاصة بعينة طبقية مكونة من ٣٣ مركزاً عمرانياً من العدد الكلي المقدر بنحو ٣٣٠ مركزاً عمرانياً يمكن تقدير متوسط عدد ورش صناعة الأحذية في كل مجموعة من المجموعات الأربع وتقدير متوسط عدد الورش في كل مركز عمراني على حدة بالإضافة إلى تقدير العدد الكلي لهذه الورش التي توجد في المراكز العمرانية لكل مجموعة على اختلاف أحجامها وفي كل منطقة موضع الدراسة، وبالتالي فإنه عن طريق استخدام مثل هذه التي تتكون من مجموعة صغيرة نسبياً من المراكز العمرانية يمكن إعطاء صورة تفصيلية عن توزيع الظاهرة قيد البحث، وتعميم نتائج التحليل الإحصائي الكمي لهذا التوزيع على جميع المراكز العمرانية في كل أرجاء المنطقة.

مثال (٢)

لدراسة تأثير طول فترة الإقامة في المنطقة التجارية لمدينة ما على مفهوم المركز التجاري للمدينة لدى السكان وتصورهم الذهني لهذا الجزء من المدينة، فإن خطة الباحث في ذلك تنحصر في إجراء سحب عينة تكون ممثلة لجميع السكان بقدر المستطاع لقياس هذا التأثير. ويحتاج الباحث بعمل إجراء عملية المعاينة أن تتوفر لديه بعض المعطيات الإحصائية من خصائص سكان تلك المنطقة التجارية والتي يمكن الحصول عليها من التعدادات السكانية أو غيرها من المصدر فمثلاً قد تتجمع لدى الباحث بيانات تفيد أن سكان المنطقة التجارية في المدينة موضع الدراسة والذين تنحصر أعمارهم بين ٣٠ و ٥٠ سنة ينقسمون من حيث مدة الإقامة في المنطقة إلى ثلاث مجموعات المجموعة الأولى تشمل السكان الذين أقاموا طوال حياتهم في المنطقة ونسبتهم ٦٠٪ من جملة السكان، والمجموعة الثانية تتكون من سكان أقاموا في المنطقة لمدة تتراوح من ٥ ـ ١٠ سنوات، ونسبتهم ٢٠٪ من جملة السكان، والمجموعة الثالثة تتضمن من أقاموا في المنطقة لمدة ٥ سنوات أو أقل ونسبتهم ٢٠٪، المجموع الكلي للسكان. وفي هذه الحالة فإن لجمع المعلومات اللازمة لمثل هذه الدراسة، يقوم الباحث بإجراء مقابلة لعينة من السكان من كل مجموعة، كأن نختار مثلاً ١٢٠ ساكناً في نفس فئة العمر (٣٠ ـ ٥٠ سنة) من الذين أقاموا في المنطقة طوال حياتهم، ٤٠ ساكناً في نفس فئة العمر أيضاً من بين الذين عاشوا في المنطقة لمدة تتراوح بين ٥ سنوات و ١٠ سنوات، ٤٠ ساكناً من الذين أقاموا لمدة ٥ سنوات أو أقل في المنطقة، وبإجراء ذلك فإن الباحث يكون قد قسم مفردات العينة على أساس مدة الإقامة إلى ثلاث طبقات يتناسب حجم مفردات كل منها مع الحجم الأصلي لنفس الطبقة من المجتمع. أما إذا أراد الباحث عمل مقارنة بين الطبقات الثلاث فإنه يقوم بأخذ ثلاث عينات فرعية Sub-Sample متساوية الحجم (٥٠ ساكناً من كل طبقة مثلاً) حتى تكون المقارنة سليمة ويعاب على الطريقة الأخيرة عند تطبيق المعاينة الطبقية أن المجموع الكلي للمفردات المختارة قد لا يكون ممثلاً Representative لجميع مفردات الظاهرة قيد البحث

و ـ العينة العشوائية المتعددة المراحل Multi-Stage Random Sample

يلائم هذا النوع من العينات العشوائية دراسة المجتمعات الكبيرة، مثل الدراسات السكانية أو الدراسات في مجال الجغرافية الاقتصادية، وهي مجتمعات يمكن تقسيمها إلى عدد من الأقسام المتشابهة التي يحتوي كل قسم منها عدد من المفردات التي تتصف بعدم التجانس في خصائصها، ولذا أطلق على هذا النوع من العينات بأنه «متعدد المراحل»، فمثلاً إذا أرادنا دراسة الحالة الاجتماعية في الريف على مستوى محافظات الجمهورية فإننا نقوم أولاً باختيار عشوائي لعدد من محافظات الجمهورية وبعد ذلك في مرحلة تالية نقوم باختيار عشوائي لعدد من مراكز المحافظات المختارة سابقاً. ثم تأتي بعد ذلك المرحلة الثالثة وفيها نقوم باختيار عشوائي لعدد من قرى المراكز المختارة في المرحلة الثانية، وتكون هذه القرى بما يختار منها عشوائياً من أسر عبارة عن المفردات التي تجري عليها الدراسة لتحديد بعض المؤثرات والمقاييس الإحصائية ويهدف التدرج السابق في أخذ العينات في مراحل إلى التبسيط والمحافظة على طبيعة المفردات غير المتجانسة داخل العينة في كل مرحلة من المراحل. ويوضح ذلك الشكل التالي (شكل رقم: ١ ـ ١).

ويعاب على المعاينة المتعددة المراحل أن كثرة عدد المراحل التي قد تتضمنها تضعف العلاقة بين معالم المجتمع الأصلي وخصائص العينة مما يؤثر بالتالي على تقدير معالم المجتمع من بيانات العينة المتحصل عليها في آخر مرحلة، كما أن هذا النوع من العينات يتطلب من الباحث جهداً كبيراً في تحديد حدود او إطار كل مرحلة وتحديد حجم العينة الفرعية المطلوبة من كل منها وذلك في ضوء الاعتبارات الخاصة بالاختيار في المعاينة العشوائية.

المرحلة الأولى — محافظات الجمهورية

١	٢	٣	٤	٥	٦	٧
٨	٩	١٠	**(١١)**	١٢	١٣	١٤
١٥	١٦	١٧	١٨	١٩	٢٠	٢١
٢٢	٢٣	٢٤	٢٥	٢٦		

المرحلة الثانية — عدد المراكز في المحافظة المختارة

١	٢	٣	**(٤)**	٥	٦
٧	٨	٩	١٠	١١	١٢

المرحلة الثالثة — عدد القرى في المركز المختار

١	٢	٣	٤	٥	٦
٧	٨	٩	١٠	١١	١٢
١٣	١٤	١٥	١٦	١٧	١٨
١٩	٢٠	٢١	**(٢٢)**	٢٣	٢٤

(شكل رقم ١ ـ ١) طريقة اختيار العينة العشوائية المتعددة المراحل

ثانياً العينات غير العشوائية (العمدية) Non-Random Samples

كثيراً ما يتعرض أسلوب المعاينة العشوائية لبعض العقبات التي تحول دون التمسك به أو الاعتماد عليه في دراسة المجتمعات وذلك عندما يتطلب سحب العينة العشوائية .إمكانيات مادية وفنية، أو عندما يجد الباحث صعوبة في الوصول إلى وحدة من وحدات المجتمع المختارة أو في حالة عدم معرفة كل مفردات المجتمع الذي ستسحب منه العينة العشوائية وفي مثل هذا الحالات يضطر الباحث إلى اتباع أسلوب التعمد والتحير الشخصي في اختيار مفردات العينة، أو ما يعرف بأسلوب المعاينة العمدية (غير الاحتمالية) وبذلك يقوم اختيار هذا النوع من العينات على أساس شخصي ولا تراعى فيه الفرص المتكافئة للمفردات لأن تكون ضمن العينة، أي لا تراعى فيه صفة العشوائية.

وكثيراً ما يستخدم المعاينة العمدية، بصفة عامة، في الأبحاث الاستطلاعية. كما في حالة تقدير معالم مجتمع كبير أو عند محاولة معرفة فكرة تقريبية سريعة عن خصائص ظاهرة ما بحيث لا تستخدم نتائجها للتعميم على المجتمع. كما تستخدم المعاينة العمدية في الاختبارات القبلية (السابقة) مثل اختبار صحيفة الأسئلة لمعرفة مدى تجاوب المبحوثين حتى يمكن إجراء التعديلات اللازمة في الأسئلة قبل بدء المعاينة الرئيسية، أو في حالة القيام ببعض القياسات لظاهرة ما لمعايرة الأجهزة المستخدمة في القياس والتأكد من سلامتها. وجدير بالذكر أنه لا توجد هناك أية طريقة إحصائية لمعرفة وقياس دقة نتائج المعاينة العمدية (غير الاحتمالية)، ولذا لا تعتبر هذه الطريقة من طرق المعاينة الجيدة إلا أنه في بعض الأحيان قد لا يجد الباحث سبيلاً ممكناً عملياً للمعاينة سوى استخدام هذه الطريقة وسنعرض فيما يلي لأهم أنواع العينات غير العشوائية (العمدية أو غير الاحتمالية وهي العينة الغرضية، العينة بالحصة، والعينة العنقودية

أ ـ العينة الغرضية

تلائم طريقة العينة الغرضية الدراسات التي تفحص الظاهرات التي تشتد فيها

درجة تباين متغايراتها، مما يجعل الباحث مصطراً إلى تحديد واختيار المتغيرات الخاصة بالبيانات المراد جمعها والتي يرى من وجهة نظره أنها تصلح للدراسة فمثلاً الباحث الذي يدرس مستوى المعيشة في الريف المصري لا يمكنه الاعتماد على الاختيار العشوائي لعينة من القرى، بل يعتمد على تحديد عدد من القرى تمثل في نظره مجتمع القرى المصرية وتكون محلاً للدراسة.

ب ـ العينة بالحصة Quata Sample

يضطر الباحث إلى استخدام مثل هذا النوع من العينات العمدية (غير الاحتمالية) عندما يتطلب منه القيام بإجراء عدد معين من المقابلات لأشخاص لهم صفات محددة في مكان معلوم أو بإجراء عدد معين من الزيارات الميدانية لجمع بيانات عن ظاهرة معينة داخل منطقة محدودة. وفي طريقة العينة بالحصة لا تختار مفردات (وحدات) العين عشوائياً ولكن تستخدم أية معلومات تساعد على الحصول على الحصة المطلوبة بسرعة وبتكاليف قليلة. ولذلك فإن هذه الطريقة تستخدم بكثرة في معاينة واستطلاع الرأي العام كما هو متبع في معهد جالوب بالولايات المتحدة الأمريكية عند التنبؤ بنتيجة الانتخابات العامة، إذ يطلب من الباحثين في هذه الحالة التعرف على رأي مجموعة من الناخبين على أن تكون من بينهم نسبة معينة من فئات مختلفة مثل أصحاب المهن الحرة وفئة العمال وفئة الموظفين... إلخ، ويترك للباحثين حرية التصرف في اختيار الأعداد المطلوبة منهم بأية طريقة يجدونها سهلة ومناسبة.

وقبل إجراء العينة بالحصة يجب التأكد من مجموعة الخصائص (ثلاث أو أربع خصائص مثلاً) التي تميز المجتمع الأصلي بحيث ترتبط هذه الخصائص ارتباطاً وثيقاً بالمتغير قيد البحث، وتصمم عينة تكون ممثلة لهذه الخصائص مجتمعة ويتضمن تصميم العينة بالحصة ثلاثة مراحل هي: ـ

١ ـ مرحلة تصنيف المجتمع الأصلي على أساس الخصائص موضع الدراسة.

٢ ـ مرحلة تحديد نسبة المجتمع في كل طبقة أوفئة

٣ ـ مرحلة تحديد الحصص التي يراد دراستها واختيارها في ضوء العدد المطلوب.

وجدير بالذكر أنه يمكن اعتبار العينة بالحصة نوع من العينات الطبقية التي يكون فيها الاختيار داخل الطبقة اختيار غير عشوائي مما قد يؤدي إلى الوقوع في خطأ التحيز من قبل الباحث من جراء التصنيف الشخص للعناصر والفئات من ناحية وعدم عشوائية الاختيار من ناحية أخرى.

جـ ـ العينة العنقودية Cluster Sample

تشبه طريقة المعاينة العنقودية العينة متعددة المراحل في كثير من مراحل اجرائها. وتقوم هذه الطريقة على أساس اختيار مفردات العينة في حزم أو عناقيد Clusters بأقل جهد وتكلفة مما يحدث بالنسبة للعينة العشوائية. فمثلاً إذا كنا بصدد دراسة مستوى المعيشة في منطقة متخلفة بأحد الأقسام الإدارية في محافظة الإسكندرية. وبافتراض عدم وجود سجل يضم سكان هذه المنطقة. ولكنهم يوجدون في سجلات مصلحة الكهرباء للمحافظة كلها بما فيها المنطقة قيد البحث، فإنه يمكن اختيار العينة على عدة تراخص أو تدريجياً بشرط أن تكون متكاملة.

ولإجراء سحب عينة عنقودية، لنفترض أن دراسة جغرافية اجتماعية (مثل دراسة خصائص النشاط السكاني والعوامل المؤثرة فيه) ستجري على سكان مدينة ما يبلغ تعدادها ٣٠٠٠٠ نسمة والمسجلين في قوائم يمكن الحصول عليها بسهولة. والمطلوب أن تكون العينة التي تجري عليها الدراسة مكونة من ٣٠٠٠ شخص فقط، ففي هذه الحالة تحتم طريقة العينة العنقودية أن تكون العينة متركزة (أو متجمعة) في أجزاء قليلة من المدينة فإذا افترضنا أن المدينة تنقسم إلى ٥٠ شياخة في كل منها ٦٠٠ شخصاً فإنه يمكن اختيار عينة من خمس شياخات فقط (أي شياخة واحدة لكل ١٠ شياخات) وتجري الدراسة على هذه الشياخات الخمس بما تحتوي من سكان

وسائل (أدوات) جمع البيانات

بعد أن يتم تحديد الأسلوب الذي على أساسه سيتم جمع البيانات سواء كان أسلوب الحصر الشامل أو أسلوب المعاينة فإنه يجب اتباع إحدى الطرق أو الوسائل (الأدوات الإحصائية) التي تستخدم في عملية الحصول على البيانات لإتمام العمل الإحصائي لأي من الأسلوبين. وأهم طرق أو أدوات جمع البيانات من مصادرها: طرق المراسلة والاتصال والعمل الحقلي (الميداني)، وتنقسم كل أداة منها إلى مجموعة من الوسائل يتم، بواسطتها اتصال الباحث بمفردات المجتمع أو العينة. وجدير بالذكر أن اختيار أي أداة من أدوات جمع البيانات يتوقف على طبيعة المعلومات التي يراد جمعها والوقت المسموح به والإمكانيات المادية المتاحة للباحث:

أولاً: المراسلة والاتصال

يعتمد الباحث على هذه الأداة في جمع البيانات إذا تعذر الوصول أو الاتصال المباشر بمفردات المجتمع أو العينة وتتم عملية جمع البيانات بهذه الأداة عن طريق إرسال استمارة البيانات الإحصائية بالبريد، أو بواسطة الاتصال التليفوني بالمبحوثين، أو حتى عن طريق نشر الأسئلة على صفحات المجلات أو الجرائد المتخصصة.

أ ـ المراسلة بالبريد ـ يقوم الباحث في هذه الطريقة بإرسال رسالة للشخص الذي وقع عليه الاختيار لاستبيانه يوضح له فيها أهمية البحث وأهدافه وسرية البيانات وعدم استخدامها إلا لغرض البحث فقط، ومع الرسالة يرفق الباحث الاستمارة الإحصائية المطلوب الإجابة على أسئلتها. كما يرسل مع الرسالة مظروف خاص بعنوان الباحث وخاص الرسوم البريدية ويطلب من المبحوث إعادة الاستمارة الإحصائية مرة أخرى إلى الباحث ومن مميزات هذه الطريقة أنها سهلة التنفيذ وقليلة التكاليف، إلى جانب أنها تعطي فرصة كافية للمبحوث في التفكير

والإجابة على الأسئلة دون ما حرج أو تردد هذا بالإضافة إلى أنها تجنب الباحث خطأ التحيز الذي قد يظهر في طريقة العمل الحقلي والاتصال المباشر بمفردات العينة أو المجتمع وعلى الرغم من ذلك ـ يعاب على طريقة المراسلة بالبريد أن نسبة الإجابات تكون عادة قليلة، وبصفة خاصة إذا كانت الاستمارة الإحصائية تحتوي على عدد كبير من الأسئلة فإن ذلك يكون سبباً في إهمال المبحوثين للاستمارات وعدم استيفائها وإعادتها للباحث. كما تتضح صعوبة هذه الطريقة في حالة إذا كان عدد من مفردات العينة يجهلون القراءة والكتابة. هذا بالإضافة إلى أن هذه الطريقة تحتاج إلى دقة بالغة في وضع الأسئلة، إذ أنه قد ينتج عن عدم فهم المبحوثين لبعض الأسئلة ووقوعهم في أخطاء تؤثر على دقة النتائج مثل خطأ التحيز في الإدلاء بالمعلومات لإنعدام الرقابة على الإجابات ولاعتقاد المبحوثين بعدم جدية أو ضرورة الدراسة.

ب ـ الاتصال التليفوني: تصلح طريقة الاتصال التليفوني كطريقة من طرق جمع البيانات، في الدراسات المحدودة التي يلعب فيها عامل الوقت دوراً مؤثراً والتي يضمن فيها الباحث وجود أجهزة تليفونية لدى المبحوثين الذين تتكون منهم العينة أو المجتمع ويمكن بهذه الطريقة مخاطبة المبحوثين والحصول منهم على الإجابات للأسئلة الموجهة إليهم.

وعلى الرغم من أن الاتصال التليفوني يعتبر من أسهل الطرق لجمع البيانات إلا أنها أصعبها من حيث إمكانية الحصول على نسبة عالية من الإجابات إذا كانت الأسئلة طويلة وتحتاج لوقت طويل في فهمها، لذا فلا بد أن تكون الأسئلة قصيرة وسريعة حتى لا تأخذ وقتاً طويلاً في الإجابة عليها.

ثانياً: العمل الحقلي (الميداني) Fieldwork

العمل الحقلي أو الميداني هو عبارة عن الفحص القريب أو التحليل في الميدان لجزء من البلاد، بما فيه من ظواهر طبيعية وبشرية، تكون سهلة

الوصول، وموضحاً مظهراً أو أكثر من مظاهر الاحتلاف السكاني وبذلك يتميز العمل الحقلي بأنه يضع الباحث وجهاً لوجه أمام مفردات ومتغيرات الظاهرة أو الظواهر المراد دراستها وتحليلها. كما أن الباحث في الميدان يستطيع أن يرى ويلمس الجوانب غير الواضحة عن الظاهرة وبالتالي يتأكد من صحة البيانات والمعلومات السابقة عنها. ولذا فإن نجاح الباحث في دراسته يتوقف إلى حد كبير على نوعية وكيفية العمل الحقلي الذي أجراه، وعلى الوقت والجهد الذي بذله، وعلى الزيارات التي قضاها في منطقة البحث

ويشمل العمل الحقلي طريقة المقابلة الشخصية (الاتصال المباشر) للمبحوثين، أو الملاحظة الميدانية وقياس الظواهر في الطبيعة، أو المسح الميداني والزيارات للمزارع والمصانع ويتوقف استخدام كل طريقة منها على خطة البحث ونوع الدراسة كما أن لكل منها مزايا وعيوب نوضحها فيما يلي:

٤ ـ المقابلة الشخصية Interviewing

تعرف المقابلة أحياناً بطريقة الاتصال المباشر لجمع البيانات إذ يتم فيها انتقال الباحث إلى المبحوثين (مفردات العينة) وذلك بفرض المواجهة الشخصية للحصول على المعلومات التي نحتاجها للدراسة كما في حالة دراسة الخصائص الاجتماعية والثقافية لسكان أحد الأقسام الإدارية في محافظة ما

وفي حالة دراسة مفردات المجتمع عددها كبير يستعين الباحث بمندوبين لجمع البيانات الذين يشترط فيهم أن يكونوا مدربين تدريباً كافياً على العمل بهذه الطريقة ويتصفون بالإضافة في تدوين البيانات.

ويمكن أن تتم المقابلة أما في أشكال محددة، أو في صورة غير محددة فهناك المقابلة المحددة أو المقفولة Closed Interview وهي المقابلة المقننة أو المهجية التي تتخذ أسلوباً منظماً حيث تكون حالة وضع الأسئلة سابقة على المقابلة نفسها، ونوجه نفس الأسئلة لجميع المفردات بدون تغير سواء في الأسلوب أو الصياغة

وهناك أيضاً المقابلة غير المحددة أو المفتوحة، وهي المقابلة غير المقننة أو غير المنهجية، التي تتميز بالأسئلة الحرة التي تتواتر بطريقة طبيعية تلقائية، أي لا تلتزم باستخدام صياغة الأسئلة وأسلوبها

ومن أهم مزايا المقابلة الشخصية أنها تلائم كثيراً دراسة المناطق التي ترتفع نسبة الأمية بين سكانها، كما أنها تتيح للباحث الحصول على معلومات أولية تقل فيها الأخطاء الصادرة من المبحوثين إلى درجة كبيرة. كذلك تعطي هذه الطريقة الفرصة لتوضيح الأسئلة الغامضة أو التي تبدو غير مفهومة المبحوثين هذا بالإضافة إلى أنه يمكن للباحث إضافة بعض الأسئلة التي يرى إضافتها أو حذف البعض الآخر تبعاً لما تمليه ظروف المقابلة كما أن الباحث يستطيع كشف أي تناقض يمكن أن يقع فيه المبحوثين.

أما عيوب طريقة المقابلة فمن أهمها احتمال تحيز الباحث أو توجيهه لمفردات المبحوثين لوجه نظر لا تخدم غرض البحث مما يؤثر على دقة النتائج، أو قد تتضمن هذه الطريقة بعض الشيء إذ من المحتمل أن يقوم الباحث باستكمال بعض الإجابات بنفسه ليتسنى له إتمام عمله في وقت قصير. وتحتاج طريقة المقابلة جهود مضنية واعتمادات مالية كبيرة إذ أنها تحتاج في بعض الأحيان إلى عدد كبير من المندوبين لجمع البيانات. كما أن هذه الطريقة لا تتمشى مع الدراسات التي تتصف بأنها تأخذ طابعاً خاصاً، بمعنى أن تكون أسئلتها محرجة أو حساسة للأفراد الذين يجري عليهم البحث والاستقصاء والذين ربما يحجمون عن الإجابة على مثل هذا النوع من الأسئلة.

ب ـ الملاحظة الميدانية وقياس الظواهر الطبيعية:

تعرف الملاحظة الميدانية بأنها المشاهدة الدقيقة لظاهرة ما، لا مع الاستعانة بأساليب البحث والدراسة لقياس وتسجيل كافة أوجه التغيرات Variations (مكانية أو زمنية) في الظاهرة وفق خطة معينة تتلائم مع طبيعة تلك الظاهرة وهي بذلك ـ الأسس والنظريات Basis and Theories التي تضبط ظاهرة أو مشكلة معينة فعند

دراسة ظاهرة أو مشكلة ما مثلاً فإننا نضع الفروض المناسبة لدراستها وحلها على أساس وضع خطة محدده تشتمل على بعض التجارب العلمية أو القياسات الحقلية، باستعمال بعض الأجهزة لقياس وتسجيل المتغيرات المتعلقة بالفروض الموضوعة، ثم نقوم باختبار صحة الفروض واحداً بعد الآخر مع استبعاد الفرض الذي لا تثبت صحته وأهميته.

وعند إجراء عملية الملاحظة فإنه يجب على الباحث مراعاة الحرص والدقة في القياس وعدم التحيز. فمثلاً عند تحليل الاختلافات المكانية أو الزمنية في شكل شاطىء منطقة ساحلية فإنه يجب قياس وتسجيل المتغيرات التي تؤثر في هذه الاختلافات حتى يمكن الوقوف على الطريقة التي تتأثر بها الظاهرة، ولرسم تحديد المناطق التي تحدث بها هذه الاختلافات أكثر من غيرها من المناطق. ومن المتغيرات التي يمكن قياسها لهذا الغرض انحدار الشاطىء Beach Slope حجم الرواسب خصائص الأمواج (طول الموجة ارتفاعها، انحدارها وقوتها، زمن الموجة واتجاهها)، خصائص التيارات الساحلية من حيث اتجاه التيار وقوته، والعوامل الجوية من ضغط جوي واتجاه الرياح وقوتها.

ويسود استخدام الملاحظة أو المشاهدة كأداة من أدوات البحث لجمع وتسجيل المعلومات في مجال الدراسات المعملية أو في التجارب الحقلية وذلك بقصد تحقيق فرض معين، أو لمعرفة علاقة بين متغيرين، أو للإيضاح لبعض النتائج التجريبية التي يكون معناها ما زال غامضاً أو مهماً، وأثناء الملاحظة يقوم الباحث بتسجيل الملاحظات أو المشاهدات في شكل قياسات معملية أو حقلية عن المتغيرات التي تحكم التجربة مثل رصد الحركة والوقت، كما في حالة رصد حركة المرور على أحد الطرق في فترات زمنية مجددة. وتجدر الإشارة هنا إلى أن الملاحظة أو المشاهدة عن طريق التجربة تمكن الباحث من السيطرة على بعض المتغيرات المحيطة بالظاهرة قيد البحث والتي لا تدخل ضمن المتغيرات التي يراد دراستها حتى يمكن اختبار سلوك الظاهرة بدقة والوقوف على حقيقة التأثير النسبي لمختلف المتغيرات ونظراً لصعوبة التحكم في جميع المحددات والعوامل

الجغرافية، وهي عوامل متشابكة، فإن الملاحظة أو المشاهدة عن طريق التجربة، الميدانية في الدراسات الجغرافية لا تتم إلا إذا كان مجال الظاهرة محدوداً والمتغيرات قيد البحث عددها قليلاً.

جـ - الزيارات

يشمل العمل الحقلي (الميداني) أيضاً جمع البيانات المنشورة التي تقيد دراسة وتحليل الظاهرة قيد البحث كالتقارير والوثائق والإحصاءات من الشركات والمؤسسات أو الهيئات الخاصة والحكومية في منطقة البحث. وكما يقول Wooldrige and East (1967) أن العمل الحقلي لا يعتبر عملاً حقيقياً إلا إذا شمل الزيارات للمزارع والمصانع ومراكز الإحصاء. ومما تجدر الإشارة إليه في هذا الصدد أنه يجب مطابقة النشرات والتقارير التي يجمعها الباحث من مصادرها على الطبيعة للوثوق من سلامتها العملية وللتأكد من صحة ما تحتويه من بيانات، كما يجب التعرف على الطرق والأساليب التي جمعت بواسطتها هذه البيانات ويتم ذلك عن طريق مناقشة المختصين وذوي الخبرة في الهيئات المسؤولة عن نشر البيانات، وإذ ما تعذر الحصول على البيانات المطلوبة أثناء القيام بالزيارات الميدانية، فإنه لا بد أن يقوم الباحث بتصميم استمارة إحصائية، أو ما يعرف «بالإستبيان» لاستكمال هذا النقص بنفسه عن طريق الاستفسار الشخصي وتدوين المعلومات عن كل أو بعض المتغيرات المطلوب دراستها.

من العجالة السابقة عن العمل الحقلي نلاحظ أن إجراءاته ووسائله تتنوع بتنوع الظروف والعوامل المتكملة في الظاهرة قيد البحث ولكن قد البحث، ولكن قد يستعين الباحث في الميدان ببعض الوسائل التي تعينه في الدراسة الميدانية بصفة عامة وفي الملاحظة بصفة خاصة وهي:

(١) تسجيل القياسات التي أجريت على المتغيرات المطلوب دراستها بواسطة الأجهزة والأدوات الخاصة، وتدوين المشاهدات الميدانية أما في جداول وأما على أجهزة التسجيل إذا تيسر استعمالها. ولتدوين القياسات والمشاهدات مكانة

خاصة في الدراسات التي توضح العلاقات بين المشاهدات والملاحظات المختلفة.

(٢) الخرائط: تعتبر الخريطة من أصلح الوسائل لمعرفة العلاقات المكانية وتفسير وتعليل كثير من غوامض الظاهرة أو الظاهرات المدروسة. ونظراً لأنه يستحيل على الباحث أن يزور كل مكان ويفحصه في الطبيعة، فلا بد له من الاعتماد على الخريطة لمعرفة الأماكن التي يصعب عليه رؤيتها في الطبيعة وعموماً فإن الباحث عن قيامه بالزيارات أو ملاحظة وقياس متغيرات ظاهرة ما في الطبيعة يحتاج إلى عدة خرائط عن الظروف الطبيعية لمنطقة الدراسة.

(٣) الصور الجوية والفوتوغرافية لظاهرات ومواقف معينة وهما من الوسائل المعينة في الدراسة الميدانية. فمن المعروف أن الباحث يرى الموقف من خلال طريقة تفكيره الخاصة واهتماماته الشخصية بالدراسة التي يقوم بها، ولكن لا يتفق باحثان على وصف واحد لظاهرة معينة أو لموقف معين أو كلا منهما يراه من وجهة نظره الخاصة.

وعموماً إذا كان منهج البحث هو الذي يحدد الطريقة المتبعة فيه، فإن الطريقة ـ بالتالي ـ هي التي تحدد أداة أو وسيلة جمع البيانات الأكثر مناسبة. إلا أن هذا لا يعني الاعتماد على أداة واحدة فقط من الأدوات السابق ذكرها لجمع البيانات لأنه يمكن جمع المعلومات المطلوب الحصول عليها. ونظراً لأن هذه الأدوات غير مستقلة تماماً عن بعضها فإن الباحث يستطيع أن يحدد الأدوات التي سيستخدمها في جمعه للبيانات والمعلومات الدقيقة التي يتعرض لها بالتحليل الإحصائي بحيث يحصل في النهاية على نتائج يتخذ على أساسها القرارات فيما بعد.

الاستمارات الاحصائية

بعد أن يتم اختيار طريقة جمع البيانات والمعلومات عن خصائص مفردات المجتمع أو العينة سواء بأسلوب الحصر الشامل أو المعاينة (العينات) يلجأ الباحث إلى عمل استمارة إحصائية خاصة، تتضمن أسئلة محددة عن تلك الخصائص المراد معرفتها وقياسها، لتكون مرشداً له في جمع بياناته ورسم إطاراً محدداً لها،

هذا فضلاً عن استخدامها كأداة لتسجيل البيانات أو قناة تستقي المعلومات من خلالها. وعادة ما تستخدم الاستمارة الإحصائية في الدراسات التي تحتاج إلى جمع بيانات كثيرة قابلة للقياس ويمكن تسجيلها بانتظام.

وكما ذكرنا آنفاً أن خطوات تصميم البحث تبدأ بوضع إطار للبيانات التي يجب الحصول عليها لاستخدامها في حل مشكلة البحث، ثم تحديد مصادر هذه البيانات والوسائل التي ستتبع في الحصول على هذه البيانات. وفي المرحلة الأخيرة ذكرنا أنه يجب أن تختار وسائل جمع البيانات اختياراً سليماً يبنى على أساس مدى ملائمة كل وسيلة، من حيث مزاياها وعيوبها لأهداف البحث وعادة ما تكون الاستمارة الإحصائية أداة هامة من أدوات أو وسائل جمع البيانات وتستخدم بالإضافة إلى الأدوات أو الوسائل الأخرى مثل المقابلة أو الملاحظة.

وهناك نوعين رئيسيين من الاستمارات الإحصائية هما: كشف البحث Schedule وصحيفة الإستبيان (أو الاستبانة) Qestionnaire ولكل منهما مزاياه وعيوبه التي نوضحها فيما يلي: -

(١) كشف البحث: يطلق اسم كشف البحث على الاستمارة الإحصائية التي تضم مجموعة من الأسئلة التي تسأل وتدون بواسطة الباحث في مقابلة شخصية للبحوث (مفردة البحث) الذي وقع عليه الاختيار في عينة البحث. كما يضم كشف البحث عند استخدامه في الملاحظة أو المشاهدة الميدانية بيان بمتغيرات الظاهرة المطلوب قياسها وجمع المعلومات عنها. ويتيح هذا النوع من الاستمارات الإحصائية للباحث درجة عالية من المرونة عند تصميمها بإعطاء الفرصة في إضافة أو حذف ما يراه الباحث من أسئلة تبعاً لظروف المقابلة الشخصية، أو صرف النظر عن بعض العوامل التي قد لا يكون لها دوراً يذكر في تباين الظاهرة موضع الدراسة. إلا أنه من أخطر عيوب كشف البحث هو احتمال تحيز الباحث لوجهة نظر شخصية لاتخذ بعرض الباحث، مما يؤثر على دقة النتائج التي ينتهي إليها البحث.

(٢) صحيفة الإستبيان: وهي عبارة عن الأداة التي تستخدم للحصول على

البيانات عن طريق الإجابة على أسئلة تتعلق بالظاهرة قيد البحث والتي يجب عليها المبحوث بنفسه، وهذه قد ترسل بالبريد أو تسلم باليد للمبحوث الذي يطلب منه في كلتا الحالتين إعادتها للباحث بعد استيفائها.

وتجدر الإشارة هنا إلى توضيح الفرق بين كشف البحث وصحيفة الإستبيان حتى لا يختلط الأمر من بينهما. فالأولى عبارة عن وسيلة قائمة بذاتها لجمع المعلومات بطريقة سريعة عن موضوعات محدد ومن مجموعة كبيرة من المفردات (المبحوثين)، بينما تستخدم الثانية كأداة لهذه الوسيلة التي يكون هدفها الأساسي ترجمة البحث العلمي إلى أسئلة معينة. وبصفة عامة تتميز صحيفة الإستيان بسهولة تنفيذها وتوفيرها للوقت والتكاليف المادية، وإتاحتها الفرصة للمبحوث في التفكير والإجابة على الأسئلة الحرجة دون تردد، بالإضافة إلى أنها تجنب الباحث الوقوع في خطأ التحيز لعدم إمكانية فرصة لرأي معين أو لوجهة نظر خاصة. إلا أن إمكانية وجود أخطاء ناجمة عن تحيز المبحوث نفسه في إجابة الأسئلة يعتبر من أهم مثالب صحيفة الإستبيان بالإضافة إلى أنها لا تصلح تماماً إذا كانت مجموعة المبحوثين في العينة أو المجتمع تحتوي على عدد كبير يجهل القراءة والكتابة أو إذا كانت البيانات المطلوبة كثيرة ووقت المبحوث ضيقاً مما يؤدي إلى تكاسل المبحوق في استيفاء الاستمارة وإعادتها للباحث.

تصميم الاستمارة الإحصائية

مهما كانت طبيعة البيانات المطلوب الحصول عليها أو الوسيلة المتبعة في جمع هذه البيانات فإنه يجب على الباحث مراعاة بعض الشروط الهامة عند تصميمه للاستمارة الإحصائية. لأن التصميم الجيد والصياغة المتقنة لأسئلة الاستمارة يعتبر أحد العوامل الجوهرية في إنجاح العمل الحقلي بصفة خاصة والبحث الذي يقوم عليه بصفة عامة، وتحتاج عملية تصميم الاستمارة الإحصائية من الباحث المعرفة الكاملة والدراية التامة بأصول صياغة الأسئلة. ورغم أن الاستمارات تختلف في تصميمها، إلا أن هناك قواعد وشروط يجب توافرها حتى يأخذ تصميم الاستمارة دوره في إنجاح البحث. هذه الشروط منها ما هو متعلق بشكل الاستمارة ومنها ما

هو يتعلق بمضمونها من حيث نوعية الأسئلة وطريقة وضعها وصياغتها.

(١) شكل الاستمارة: لا شك أن الاهتمام بشكل الاستمارة الإحصائية يعتبر من العوامل الرئيسية في عملية جمع البيانات الدقيقة غير المشكوك فيها، حيث يشجع الشكل الجيد المبحوثين على الاستجابة لمحتواها. ويتحدد شكل الاستمارة الجيد بعدة عوامل منها:

أ - جودة الاستمارة من حيث نوع الورق المستخدم الذي يجب أن يكون من النوع يتحمل الاستخدام الكثير في تدوين المعلومات.

ب - حجم الاستمارة من حيث صفحات الاستمارة التي يجب أن لا تكون قليلة على حساب الأماكن الخالية المخصصة للإجابة أو لا تكون كثيرة حتى لا يكون ذلك سبباً في إرهاق المبحوثين في الإجابة على أسئلتها.

جـ - ترتيب وتنظيم الأسئلة داخل الاستمارة، إذ أن التسلسل والترتيب في وضع الأسئلة (عن طريق إعطاء الأسئلة أرقاماً تدريجية أو وضع الأسئلة في شكل مجموعات أو تقسيمات متجانسة تترابط فيما بينها ترابطاً منهجياً يمكن معه حمس المطلوب، بحيث يبدأ من الأسئلة البسيطة إلى الأسئلة المركبة، أو من أسئلة تتميز بالشمول إلى أسئلة تتميز بالتركيز على أفكار دقيقة ومحددة) يعتبر من أهم الشروط التي يجب مراعاتها عند تصميم الاستمارة الإحصائية مهما كان نوعها لأن ذلك يساعد على سهولة الإجابة، كما يعمل على تسهيل عملية التحليل والدراسة بعد ذلك.

وعموماً يجب أن يظهر عنوان البحث بوضوح في صدر الاستمارة وكذلك باسم الهيئة أو الجهة المشرفة على الدراسة، الإضافة ما يشير إلى سرية استخدام بيانات الاستمارة إلا لعرض البحث فقط، مع وضع بعض التعليمات المختصرة والمبسطة لتوضيح أهداف الدراسة إن أمكن ذلك. ونظراً لأن معظم التحليلات الإحصائية تقوم بها في الوقت الحاضر أجهزة الحاسب الآلي فمن المستحسن أن تتضمن الاستمارة رموزاً Codes حتى تسهل مهمة نقلها وتفريغها على البطاقات الخاصة بالحاسب الآلي.

(٢) مضمون الاستمارة: يقصد بمضمون الاستمارة هو كيفية صياغة الأسئلة التي تعتبر ذات أهمية بالغة في الحصول على إجابات صحيحة وبالتالي على معلومات دقيقة. وكلما كانت الأسئلة أو التعبير عما هو مطلوب، واضح دون ما صعوبة أو تعقيد لفظي أو سوء فهم كلما سهلت مهمة الباحث والمبحوث في نفس الوقت. وبصفة عامة فإنه يمكن تحقيق ذلك بأن تكون الأسئلة على شكل حوار طبيعي تلقائي، أي ليس المقصود بها أن نتوصل إلى إجابات معينة، مع تجنب الأسئلة الطويلة التي تزيد من احتمالات سوء الفهم كما يجب أن تكون الأسئلة محددة ودقيقة حتى نحصل على معلومات صحيحة، أي يجب أن يعطي كل سؤال فكرة واحدة واضحة عما يطلب السؤال عنه. فمثلاً يبدو السؤال أين كان ميلادك؟ غامضاً. والأفضل منه يكون السؤال، في أي قرية أو مدينة كان ميلادك؟ وهو يبدو أكثر تحديداً ووضوحاً من السؤال الأول. كما يجب أن تكون الأسئلة بعيدة تماماً عن الأسئلة الحرجة ذات الحساسية البالغة. ويستطيع الباحث التحايل على ذلك بصياغة أسئلة غير مباشرة، فمثلاً يمكن التعرف على مقدرة ودخل العامل بطرح الأسئلة التي تستفسر عن طبيعة العمل الذي يقوم به العامل. على أنه يجب أن يكون الباحث لبقاً وذكياً عند وضع الأسئلة حتى لا يضع الأسئلة توحي بإجابات معينة أو أسئلة افتراضية تكون الإجابة عليها غير مفيدة. فمثلاً يمكن طرح السؤال: ما مقدار الأجر الإضافي الذي ترغب أن تحصل عليه شهرياً حتى يتحسن مستوى معيشتك؟ بدلاً من السؤال: هل تكون راضياً لو ارتفع مرتكب الشهري إلى ٦٠ جنيهاً؟ كذلك يجب وضع تفسيرات محددة للمصطلحات التي تكون مجالاً للشك من حيث الفهم وتوضيحات دقيقة للتعريفات المستخدمة مثل تعريف الأسرة أو الدخل. كما يجب أن تصاغ الأسئلة أما لتوضيح الآراء والاتجاهات Attiudes أو لتوضيح الحقائق مثل السن والمهنة أو الملكية الزراعية أو العقارية.

وقبل إتمام صياغة الاستمارة الإحصائية، ينبغي على الباحث أن يتفهم طبيعة المبحوثين موضع الدراسة وذلك عن طريق تصميم استمارة استطعية Piolt Questionnare توزع على عينة ذات عدد محدود من الأفراد ليست لهم علاقة

بالبحث ليجيبوا على أسئلتها. ومن طريقة الإجابة في الاستمارة الاستطلاعية يمكن التعرف على الأسئلة التي يمكن أن تكون غامضة أو غير مفهومة لتعاد صياغتها بعد توضيحها، كما تجدر الإشارة في هذا الصدد إلى أنه يجب على الباحث أن يضع بعض الأسئلة للمراجعة Checking Questions للتأكد من صحة الإجابات خصوصاً عند وجود تعارض في الإجابات على هذا النوع من الأسئلة وإجابات الأسئلة الخاصة في الاستمارة والتي تحمل نفس الإجابة أو عكسها فمثلاً السؤال هل تحب عملك؟ يتعارض مع السؤال: هل تتغيب كثيراً عن العمل؟ فإذا كانت الإجابة على السؤال الأول بالإيجاب وعلى الثاني بالنفي فإن ذلك يؤكد أن حب العمل لا يؤدي إلى التغيب كثيراً عن العمل.

عرض البيانات عن طريق الأشكال البيانية

أوضحنا فى مقدمة هذا الباب كيفية انشاء الجداول للبيانات النوعية والكمية وتأكد أن هذه الجداول تعتبر أداة قيمة فى دراسة العلاقات المتداخلة بين اثنين أو أكثر من العناصر أو المتغيرات .

ولاشك أن الاشكال البيانية تعتبر من أكثر الأساليب والأدوات الشائعة فى عرض البيانات الاحصائية وخاصة إذا تم توجيه الرسالة المطلوب عرضها للقارئ العادى أو تقديمها لنوعية من القراء الذين ليس لديهم الوقت الكافى لتفهم مدلولات الأرقام أو تفهم العلاقات بين العناصر المختلفة وخاصة فى حالة الجداول التى تحتوى على بيانات كثيرة . ولذلك يتم استخدام الاشكال البيانية باعتبار أن القارئ عادة يجتذب بسرعة للشكل البيانى عن الجداول . ومن المهم ان نؤكد أنه ليس هناك قواعد ثابتة لاختيار أى الاشكال البيانية التى يجب رسمها او أن هناك اسلوب بيانى محدد الذى يعتبر صحيحا لموقف معين ولكن يمكن القول أن الاشكال البيانية يمكن تقييمها فى ضوء الفعالية فى عرض البيانات التى تسمح للقارئ أو لمستخدم البيانات فى الحصول على أهم المفاهيم من الشكل البيانى بسهولة ويسر ودقة . كما يجب أن نتذكر دائما أنه من الأفضل انشاء مجموعة أو سلسلة من الاشكال البيانية توضح كل منها مفهوما واحدا لفاعلية بدلا من اعداد شكل بيانى واحد يتضمن كثير من العلاقات التى قد يصعب دراستها بفعالية . وهناك أشكال بيانية للمتغيرات النوعية والكمية منها :

الاعمدة الرأسية والأفقية

يعتبر هذا الاسلوب من أشهر اساليب العرض البيانى للبيانات بصورة فعالة سواء بالنسبة للبيانات النوعية أو الكمية .

وتستخدم الأعمدة الرأسية أو الافقية (مستطيلات) لعرض البيانات ويطلق على مجموعة الأعمدة الرأسية التى تكون اطوالها متناسبة مع القيم التى نرغب فى عرضها بشكل الأعمدة الرأسية . وكذلك الحال يشار الى مجموعة الأعمدة الأفقية التى يكون اطوالها أيضا متناسبة مع القيم التى نرغب فى عرضها بشكل الأعمدة الأفقية .

وحتى تكون الاشكال متوازنة فانه يجب اختيار قواعد متساوية لجميع الأعمدة سواء كانت رأسية أو افقية حتى تتناسب مساحات هذه الأعمدة مع ارتفاعاتها باعتبار ان مساحة العمود تساوى قاعدة العمود فى ارتفاعه . وكذلك يجب ان يبدأ قياس ارتفاع الأعمدة من الصفر حتى يمكن الاحتفاظ بهذه النسب .

ويلاحظ أنه إذا كان هناك متغير واحد نرغب في عرضه بواسطة اسلوب الأعمدة لعدد من السنوات أو الأماكن ... الخ فانه يجب رسم العمود لكل سنة أو مكان أما إذا كان هناك متغيرين أو أكثر لعدد من السنوات أو الأماكن فانه يجب رسم عمودين أو أكثر لكل سنة أو مكان باستخدام التظليل الخفيف والتظليل الداكن او استخدام الألوان منعا من الغموض أو الالتباس عند النظر الى الشكل البيانى وقد يفضل البعض رسم أكثر من متغير فى عمود واحد مع تقسيمه الى أجزاء تتناسب مع قيمة كل متغير باستخدام التظليل الخفيف والتظليل الداكن وكذلك الألوان المختلفة . أما إذا كان المتغير الذى نرغب فى عرضه يمكن ان يأخذ قيما سالبة كمتغير الربح والخسائر فانه يمكن عرض الشكل البيانى عن طريق تمثيل الأرباح بأعمدة فوق المحور الأفقى للشكل البيانى وتمثيل الخسائر باعمدة اسفل المحور الأفقى للشكل البيانى .

ومن الملاحظ أنه يجب الأخذ فى الاعتبار عدة أمور أخرى عند رسم الأعمدة الرأسية والافقية أن نكتب أسفل أو بجانب كل عمود المتغير الزمنى أو النوعى الخاص بكل عمود مع كتابة عنوان مختصر للشكل البيانى . كما أنه فى حالة وجود عمود طويل جدا ضمن مجموعة الأعمدة وسيؤثر على الشكل الجمالى للرسم فانه يمكن استثناء قطع أو كسر هذا العمود الطويل هذا دون المساس بالإتجاه العام للأعمدة الأخرى . كما يجب أيضا مراعاة ترك مسافة بين كل عمود والعمود التالى له ومع ذلك يمكن اهمال ذلك حتى تعطى تاثيرا مستمرا كما هو الحال بالنسبة للتوزيعات التكرارية التى سبق الإشارة اليها

مثـــــال :

يمثل الجدول التالى توزيع عدد السائحين بالآلاف من عام ١٩٩٢ حتى عام ١٩٩٧ حسب الجنسية.

السنة	عدد السائحين بالألف			إجمالى
	عرب	دول صناعية متقدمة	اخرون	
١٩٩٢	١١٠٢,٩	١٨٦٦,٠	٢٣٨,٠	٣٢٠٦,٩
١٩٩٣	٩٢٢,٤	١٢١١,٦	٣٧٣,٨	٢٥٠٧,٨
١٩٩٤	٩٣١,٧	١١٣٦,٦	٥١٣,٧	٢٥٨٢,٠
١٩٩٥	٨٢٢,٩	١٥٩٩,٦	٧١٠,٩	٣١٣٣,٤
١٩٩٦	٨٩٦,٢	٢١٣٣,٢	٨٦٦,٥	٣٨٩٥,٩
١٩٩٧	٩٦٦,٨	٢٢٢٦,٢	٧٦٨,٤	٣٩٦١,٤

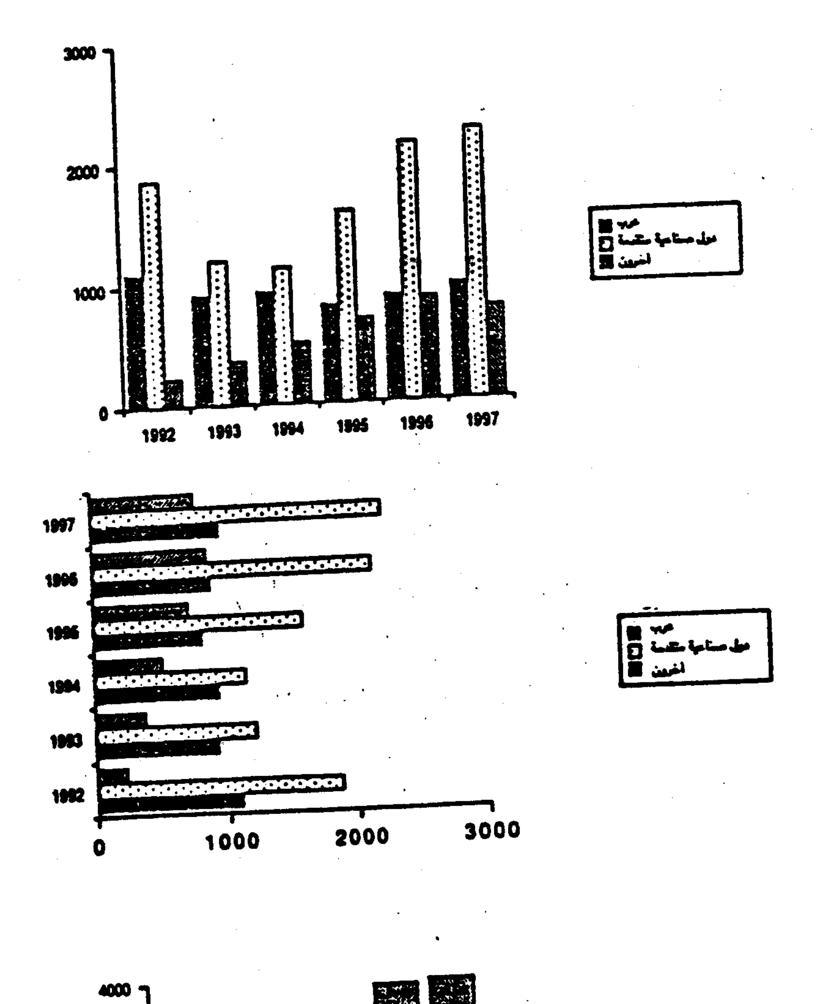
المصدر : نشرة البحوث السياحية : وزارة السياحة

والمطلوب تصوير هذه البيانات باستخدام أسلوب الأعمدة رأسيا وأفقيا على أن تمثل نوعية السائحين للسنة الواحدة بثلاث أعمدة متلاصقة مرة وعلى أن تمثل نوعية السائحين فى عمود واحدة مرة أخرى.

الحـــل :

١ - نرسم محورين متعامدين المحور الأفقى والمحور الرأسى ويخصص المحور الأفقى للسنوات والمحور الرأسى لعدد السائحين.

٢ - نختار مقياس رسم حسابى معين يتناسب مع حجم الورقة التى سيتم عليها الرسم والمساحة المخصصة للرسم حتى لا يكون الرسم صغيرا جدا أو كبيرا جدا.


3000
2000
1000
0
1992 1993 1994 1995 1996 1997
عرب
دول صناعية متقدمة
آخرون
1997
1996
1995
1994
1993
1992
0 1000 2000 3000
عرب
دول صناعية متقدمة
آخرون
4000
3000
2000
1000
0
1992 1993 1994 1995 1996 1997
عرب
دول صناعية متقدمة
آخرون

لوحــة الدائــرة :

تمثل الدائرة أحد الاشكال البيانية الشائعة الاستخدام والمفضلة بين الاحصائيين وتستخدم الدائرة لإيضاح الأحجام النسبية للمكونات داخل التجمع الكلى أو الاجمالى مع الأخذ فى الاعتبار أن يكون عدد هذه المكونات ليس كبيرا . ولذلك فان العرض البيانى أو التصويرى لتقسيمات أحد المتغيرات بحيث تكون مساحة هذه القطاعات متناسبة مع أحجام التقسيمات تحت الدراسة يطلق عليه لوحة الدائرة .

وتستخدم الدائرة بمعرفة الباحثين أو الشركات والدوائر الرسمية والصحف والمجلات لوصف الموازنات او لوصف توزيع التكاليف فى شركة معينة أو فى عرض نتائج انتخابات أو عرض نتائج استقصاء للرأى العام أو فى عرض نوعيات متغير معين مثل توزيع عدد السائحين حسب الجنسية الى غير ذلك .

ولذلك فان أول خطوة لإعداد لوحة الدائرة هى رسم دائرة : والدائرة تمثل العدد الكلى للمشاهدات نقوم بعد ذلك بتحويل المشاهدات الى تكرارات نسبية (أو نسب مئوية من المجموع الكلى لقيمة المتغير) ثم نقوم بتجزئة الدائرة (التى قمنا برسمها) الى عدد من الشرائح بحيث تمثل كل شريحة نسق معين وأن يكون حجم الشريحة متناسبا مع التكرار النسبى (النسبة المئوية) لكل نسق معين . وحيث ان الدائرة لها ٣٦٠ درجة فنقوم بضرب التكرار النسبى (النسبة المئوية) لكل نسق فى عدد درجات الدائرة وهى ٣٦٠ درجة لنحصل على زاوية كل شريحة فى الدائرة التى تمثل كل نسق

مثــــال :

افترض أن لديك البيانات التالية عن عدد السائحين فى عام ١٩٩٦، ١٩٩٧ موزعين حسب الجنسيات كالآتى :

السنة	عدد السائحين بالألف				إجمالى
	عرب	دول صناعية متقدمة	دول شرق أوروبا والصين	اخرون	
١٩٩٦	٨٩٦,٢	٢١٣٣,٢	٢١٥	٦٥١,٥	٣٨٩٥,٩
١٩٩٧	٩٦٦,٨	٢٢٢٦,٢	١٩٨,٩	٥٦٩,٥	٣٩٦١,٤

المصدر : نشرة البحوث السياحية : وزارة السياحة

والمطلوب تمثيل بيانات كل سنة بدائرة.

الحـــل :

بالنسبة لعام ١٩٩٦

١ - نقوم بتحويل المشاهدات (عدد السائحين لكل نوع) إلى تكرارات نسبية (نسب من المجموع الكلى للسائحين) كالتالى :

التكرار النسبى لعدد السائحين العرب = ٢ر٨٩٦ ÷ ٩ر٣٨٩٥ = ٢٣٠ر٠

التكرار النسبى لعدد السائحين من الدول الصناعية المتقدمة

= ٢ر٢١٣٣ ÷ ٩ر٣٨٩٥ = ٥٤٨ر٠

التكرار النسبى لعدد السائحين من دول شرق أوروبا والصين

= ٢١٥ ÷ ٩ر٣٨٩٥ = ٠٥٥ر٠

التكرار النسبى لعدد السائحين المصنفين آخرون

= ٥ر٦٥١ ÷ ٩ر٣٨٩٥ = ١٦٧ر٠

مجموع التكرات النسبية = ٠٠ر١

٢ - نحدد زاويا كل نسق (نوعية السائحين) فى الدائرة عن طريق ضرب التكرار النسبى لكل نسق فى مجموع درجات الدائرة وهى ٣٦٠ درجة.

زاوية السائحين العرب = ٢٣٠ر٠ × ٣٦٠ = ٨ر٨٢ درجة فى الدائرة

زاوية سائحى الدول المتقدمة صناعيا = ٥٤٨ر٠ × ٣٦٠ = ٢٨ر١٩٧ درجة فى الدائرة

زاوية السائحين فى دول شرق أوروبا والصين = ٠٥٥ر٠ × ٣٦٠ = ٨ر١٩ درجة فى الدائرة

زاوية السائحين الآخرين = ١٦٧ر٠ × ٣٦٠ = ١٢ر٦٠ درجة فى الدائرة.

٣ - نقوم بعد ذلك برسم دائرة نقوم بتجزئتها حسب زوايا كل منها من أنواع السائحين.

بالنسبة لعام ١٩٩٧

نقوم بنفس الخطوات السابقة التى أجريناها فى عام ١٩٩٦

١ - تحويل المشاهدات (عدد السائحين لكل نوع) إلى تكرارات نسبية (نسب من المجموع الكلى للسائحين) كالتالي :

التكرار النسبى لعدد السائحين العرب = ٩٦٦ر٨ ÷ ٣٩٦١ر٤ = ٢٤٤ر٠

التكرار النسبى لعدد السائحين من الدول الصناعية المتقدمة

= ٢٢٢٦ر٢ ÷ ٣٩٦١ر٤ = ٥٦٢ر٠

التكرار النسبى لعدد السائحين من دول شرق أوروبا والصين

= ١٩٨ر٩ ÷ ٣٩٦١ر٤ = ٠٥٠ر٠

التكرار النسبى لعدد السائحين المصنفين آخرون

= ٥٦٩ر٥ ÷ ٣٩٦١ر٤ = ١٤٤ر٠

مجموع التكرات النسبية = ٠٠ر١

٢ - نحدد زاويا كل نسق (نوعية السائحين) فى الدائرة عن طريق ضرب التكرار النسبى لكل نسق فى مجموع درجات الدائرة وهى ٣٦٠ درجة.

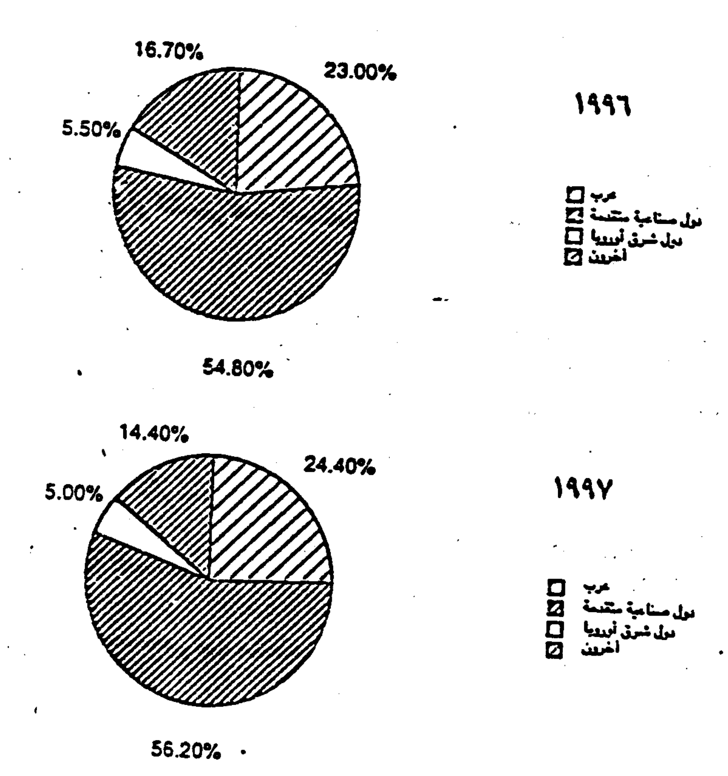
زاوية السائحين العرب = ٢٤٤ر٠ × ٣٦٠ = ٨٤ر٨٧ درجة فى الدائرة

زاوية سائحى دول متقدمة = ٥٦٢ر٠ × ٣٦٠ = ٣٢ر٢٠٢ درجة فى الدائرة

زاوية السائحين من دول شرق أوروبا والصين = ٠٥٠ر٠ × ٣٦٠ = ١٨ درجة فى الدائرة

زاوية السائحين المصنفين آخرون = ١٤٤ر٠ × ٣٦٠ = ٨٤ر٥١ درجة فى الدائرة.

٣ - نقوم بعد ذلك برسم دائرة نقوم بتجزئتها حسب زوايا كل منها من أنواع السائحين.



اشكال الخطوط البيانية

عندما يكون لدينا سلسلة زمنية تتضمن عدد كبير من الفترات الزمنية فانه من المفضل استخدام خط بيانى يوضح التقلبات الواضحة فى السلسلة الزمنية أو فى حالة وجود عدة سلاسل زمنية مطلوب ايضاحها بيانيا فى شكل واحد .

ولاشك ان بداية رسم الخط البيانى لسلسلة زمنية واحدة أو الخطوط البيانية لعدة سلاسل زمنية هى أن نرسم محورين متعامدين أحدهما يمثل المحور الأفقى ويخصص لمتغير الزمن والمحور الرأسى يخصص للمتغير تحت الدراسة . ويلاحظ أنه يجب تقسيم المحور الأفقى الى أقسام متساوية بقدر عدد السنوات واذا كانت هناك سنة فى وسط البيانات ليس بها بيانات فلانهملها كما يقسم المحور الرأسى الى أقسام متساوية بحيث يتدرج فى الكبر الى أن يحتوى على أكبر قيمة للمتغير تحت الدراسة كما يجب أن يتدرج المحور الرأسى بدءا من الصفر . ويقوم الباحث بعد ذلك بتمثيل كل مشاهدة للمتغير لكل فترة زمنية بنقطة يتم تحديدها على المحور الرأسى طبقا لمقياس الرسم المختار . نقوم بعد ذلك بوصل النقط المتتالية بخطوط مستقيمة .

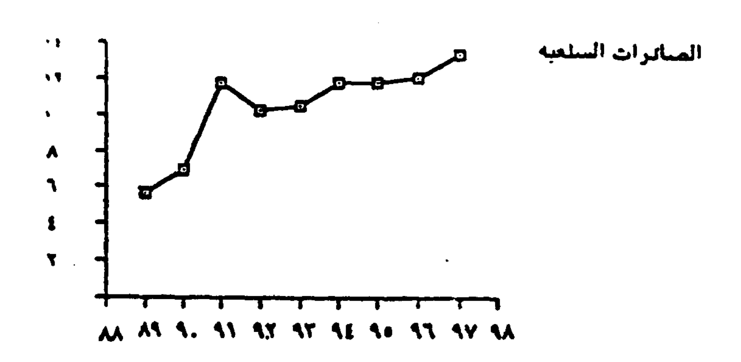
مثـــال :

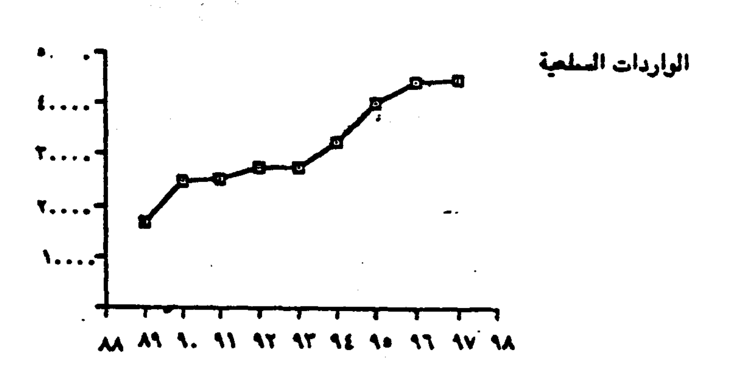
يمثل الجدول التالى صادرات وواردات مصر السلعية وعجز الميزان التجارى بملايين الجنيهات فى الفترة ١٩٨٩ حتى ١٩٩٧.

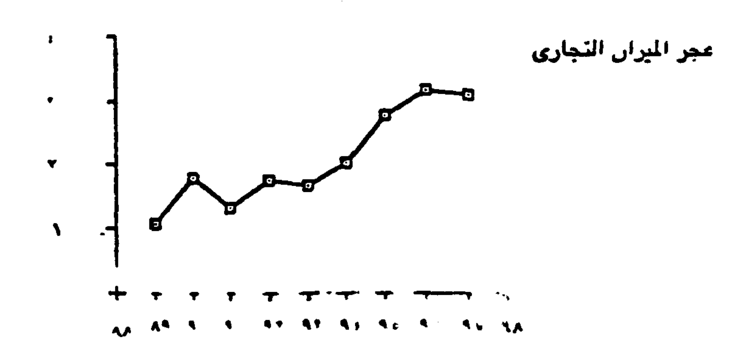
السنة	الصادرات السلعية	الواردات السلعية	عجز الميزان التجارى
١٩٨٩	٥٧٣٤.٧	١٦٦٣٢.٦	١٠٨٨٨.٩
١٩٩٠	٦٩٥٣.٨	٢٤٨٢٣.٢	١٧٨٦٩.٤
١٩٩١	١١٧٦٤.٧	٢٥٢١٦.٣	١٣٤٥١.٦
١٩٩٢	١٠١٧١.٢	٢٧٦٥٦.١	١٧٤٨٤.٩
١٩٩٣	١٠٤٦٤.٥	٢٧٥٥٠٠.٤	١٧٠٨٥.٩
١٩٩٤	١١٧٥٧.٥	٣٢٤٦٠٠.٦	٢٠٧٠٣.١
١٩٩٥	١١٧٠٣.٨	٣٩٨٩٠٠.٩	٢٨١٨٧.١
١٩٩٦	١٢٠٠٤.١	٤٤٢١٧.٩	٣٢٢١٣.٨
١٩٩٧	١٣٢٨٥.٩	٤٤٧٦٨.٨	٣١٤٨٢.٩

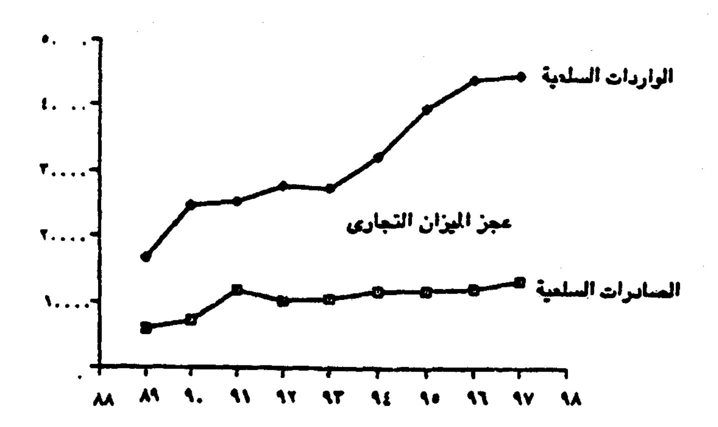
المصدر : الجهاز المركزى للتعبئة العامة والاحصاء.

والمطلوب تصوير الصادرات والواردات وعجز الميزان التجارى بخط بيانى لكل منها فى شكل بيانى منفصل ثم رسم هذه المتغيرات بثلاث خطوط بيانية فى شكل واحد ثم رسم الصادرات والواردات بخطين بيانيين فى شكل واحد واستخراج العجز التجارى بالرسم


الصادرات السلعيه
١٢
١٠
٨
٦
٤
٢
٨٨ ٨٩ ٩٠ ٩١ ٩٢ ٩٣ ٩٤ ٩٥ ٩٦ ٩٧ ٩٨


الواردات السلعية
٤٠٠٠٠
٣٠٠٠٠
٢٠٠٠٠
١٠٠٠٠
٨٨ ٨٩ ٩٠ ٩١ ٩٢ ٩٣ ٩٤ ٩٥ ٩٦ ٩٧ ٩٨


عجز الميزان التجارى
١
٩٨



الاشكال البيانية للتوزيعات التكراوية

لاشك أن الاشكال البيانية تعتبر من الأساليب المفضلة لعرض البيانات الاحصائية إذ أن البيانات يجرى اختصارها أو خفضها فى أى صورة من الجداول السابق الاشارة اليها ومن الطبيعى يمكن عرض هذه الجداول بيانيا . وقد اوضحنا سلفا كيفية عرض البيانات الموضوعة فى صورة جداول بيانيا : إلا أنه يتبقى كيفية عرض جدول التوزيعات التكرارية بيانيا وهو ما سنعرضه فى الصفحات القادمة

ولاشك أن أول الاشكال البيانية الشائعة الاستخدام فى تصوير التوزيعات التكرارية هو ما يطلق عليها المدرج التكرارى

أولا : المدرج التكرارى HISTOGRAM

يعرف المدرج التكرارى بأنه شكل بيانى للتوزيع التكرارى المطلق أو التوزيع التكرارى النسبى لمجموعة من البيانات الكمية المتصلة فى صورة مستطيلات متلاصقة تتناسب مساحتها مع تكرارات الفئات وما دامت الفئات متساوية فان ارتفاع المستطيلات تتناسب مع التكرارات .

إذ أن مساحة مستطيل الفئة الأولى (القاعدة × الارتفاع) مقسومة على تكرار الفئة الثانية وهكذا . ولذلك فى حالة أن يكون التوزيع التكرارى متساوى الفئات ستجد النسبة بين مساحة مستطيل كل فئة مقسومة على تكرار الفئة ستنتهى الى النسبة بين ارتفاع المستطيل لكل فئة وتكرار هذه الفئة .

ولرسم المدرج التكرارى نرسم محورين متعامدين ثم نضع الفئات على المحور الأفقى والتكرارات على المحور الرأسى .

ويلاحظ أنه يجب تقسيم كل من المحورين الى تقسيمات متساوية ويغطى المحور الرأسى أكبر تكرار موضح بالتوزيع التكرارى مع ملاحظة أن يبدأ المحور الرأسى بالصفر .

أما المحور الأفقى فلا يشترط أن يبدأ بالصفر وانما يحتوى فقط على الحدود الدنيا للفئات وفئة أخرى اضافية فى النهاية نقيم بعد ذلك على طول كل فئة مستطيلا يتناسب مع تكرارها أو نقيم أعمدة من الحدود الدنيا للفئات ثم نصل هذه الأعمدة بخطوط افقية ويطلق على الشكل الناتج المدرج التكرارى ومساحته الكلية تمثل مجموع التكرارات

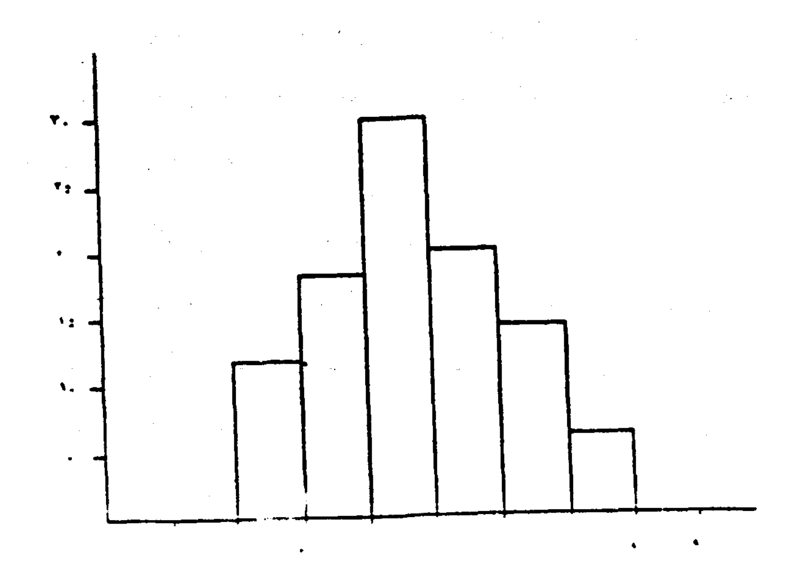
مثال

اذا كان لديك التوزيع التكرارى التالى لمبيعات ١٠٠ شركة بالاف الجنيهات فالمطلوب رسم المدرج التكرارى للتكرارات المطلقة

فئات (المبيعات)	التكرار المطلق (عدد الشركات) فى الفئة
٢٠ –	١٢
٣٠ –	١٨
٤٠ –	٣٠
٥٠ –	٢٠
٦٠ –	١٤
٧٠ –	٦
المجموع	١٠٠

الحـــل:

المدرج التكرارى للتكرارات المطلقة



المدرج التكرارى فى حالة الفئات غير المتساوية

اوضحنا سلفا كيفية رسم المدرج التكرارى فى حالتى التكرارات المطلقة والنسبية بافتراض تساوى فئات التوزيع التكرارى . كما سبق التأكيد والايضاح أن التكرارات تمثلها المساحات وقد استخدمت ارتفاعات المستطيلات لأن قواعدها كانت متساوية ومن ثم تتناسب المساحات مع الارتفاعات

أما إذا كانت الفئات غير متساوية فان معنى ذلك اختلاف قواعد المستطيلات وحيث اننا نرغب دائما أن تكون المساحات هى التى تعبر عن التكرارات ، لذلك يجب حساب ارتفاع المستطيلات حتى يسهل رسمها ويمكن اجراء ذلك عن طريق حساب كثافة التكرار المطلق أى قسمة تكرار كل فئة على طول الفئة وهى ما يجب قياسها على المحور الرأسى فى حالة ما إذا كانت مساحات مستطيلات المدرج التكرارى تناظر تكرارات الفئات ومن ثم نقيم على طول كل فئة مستطيلا يكون ارتفاعه عبارة عن كثافة التكرار وبذلك تكون مساحات هذه المستطيلات عبارة عن تكراراتها .

مثـــال :

يوضح الجدول التالى التوزيع التكرارى لمبيعات ١٠٠ شركة بالآف الجنيهات والمطلوب رسم المدرج التكرارى للتكرارت المطلقة .

فئـــات (المبيعـــات)	التكرار المطلق (عدد الشركات) فى الفئة
٢٠ –	١٢
٣٠ –	١٨
٥٠ –	٣٠
٦٠ –	٢٠
٧٠ –	١٤
٨٠ – ١٢٠	٦
المجموع	١٠٠

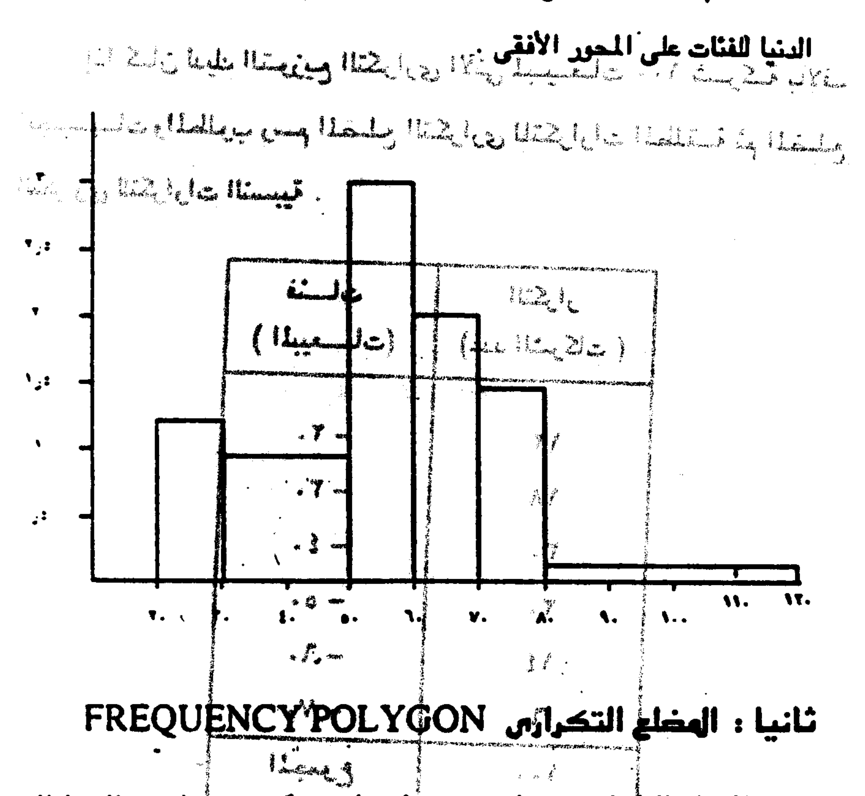
الحـــل :

المدرج التكرارى المطلق

يلاحظ ان فئات التوزيع التكرارى غير متساوية ومن ثم يجب ايجاد كثافة التكرار المطلق :

فئات (المبيعات)	التكرار المطلق (عدد الشركات)	طول الفئة (عددالوحدات من ١٠ آلاف جنيه)	كثافة التكرار المطلق (عدد الشركات لكل ١٠ آلاف جنيه)
٢٠ –	١٢	١٠	١٫٢٠
٣٠ –	١٨	٢٠	٫٩٠
٥٠ –	٣٠	١٠	٣٫٠
٦٠ –	٢٠	١٠	٢٫٠
٧٠ –	١٤	١٠	١٫٤٠
٨٠ – ١٢٠	٦	٤٠	٫١٥
المجموع	المساحة = القاعدة × الارتفاع		

نقوم بعد ذلك بوضع كثافة التكرار المطلق على المحور الرأسى والحدود الدنيا للفئات على المحور الأفقى



ثانيا : المضلع التكرارى FREQUENCY POLYGON

المضلع التكرارى عبارة عن خط بيانى منكسر يصل بين النقط التى رصدت طبقا لمراكز الفئات على المحور الأفقى والتكرارات المطلقة أو النسبية على المحور الرأسى مع وجود فئة يكون تكرارها صفرا عند بداية ونهاية الخط البيانى وذلك لقفل الشكل الا انه يمكن وضع الحدود الدنيا للفئات على المحور الأفقى ولكن عند رصد النقط يتم ذلك أمام مراكز الفئات ويستخدم هذا الاسلوب عندما يكون التوزيع التكرارى متساوى الفئات .

مثال :

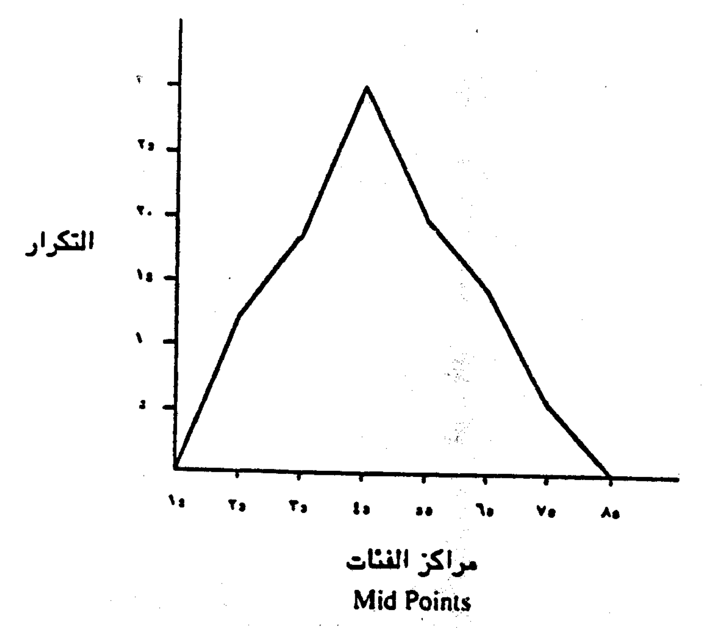
إذا كان لديك التوزيع التكرارى الآتى لمبيعات ١٠٠ شركة بآلاف الجنيهات والمطلوب رسم المضلع التكرارى للتكرارات المطلقة ثم المضلع التكرارى للتكرارات النسبية .

فئات (المبيعات)	التكرار (عدد الشركات)
٢٠ –	١٢
٣٠ –	١٨
٤٠ –	٣٠
٥٠ –	٢٠
٦٠ –	١٤
٧٠ –	٦
المجموع	١٠٠

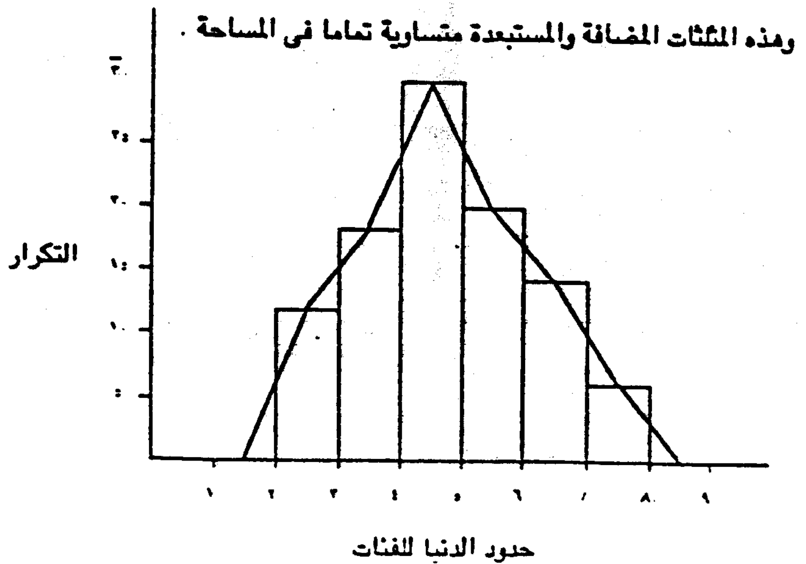
الحــل :

المضلع التكرارى للتكرارات المطلقة

نرسم محورين متعامدين ونضع مراكز الفئات على المحور الأفقى وهى عبارة عن ٢٥ ، ٣٥ ، ٤٥ ، ٥٥ ، ٦٥ ، ٧٥ ، ٨٥ والتكرار المطلق على المحور الرأسى ثم نرصد النقط ونصلها بخطوط مستقيمة



ويلاحظ أنه يمكن رسم المضلع التكرارى من المدرج التكرارى وذلك عن طريق توصيل منتصف قمم المستطيلات المتتالية ويدخل فى ذلك الفئات التى يكون تكرارها يساوى صفرا . ويلاحظ أن المساحة تحت الخط البيانى المقفل يساوى مساحة المدرج التكرارى وذلك لأنه اضيفت مثلثات على المستطيلات وهذه المثلثات المضافة والمستبعدة متساوية تماما فى المساحة .



و نود أن نشير أن القواعد السابقة لانشاء المضلع التكرارى سواء للتكرارات المطلقة او النسبية تعتبر صحيحة فقط عند تساوى الفئات فى التوزيع التكرارى أما إذا كانت الفئات غير متساوية وجب ايجاد كثافة التكرار ووضعها على المحور الرأسى ومراكز الفئات على المحور الأفقى لرسم المضلع التكرارى.

ثم نقوم بعد ذلك برصد النقط طبقا لمراكز الفئات وكثافة التكرارات المناظرة لها ثم ايصال النقط المتتالية بخطوط مستقيمة مع ملاحظة اقفال الشكل .

ويلاحظ انه يفضل استخدام المدرج التكرارى عند وجود عدد قليل من الفئات واستخدام المضلع التكرارى عند وجود عدد كبير من الفئات . ويستخدم ايهما فى الأحوال العادية . ومن غير شك تفيد التوزيعات التكرارية فى تلخيص البيانات وعرضها بصورة سهلة وسريعة .

ثالثا : المنحنى التكرارى FREQUENCY CURVE

يمكن تقريب شكل المدرج التكرارى الى منحنى تكرارى ممهد وكذلك الحال يمكن تقريب شكل المضلع التكرارى الى منحنى تكرارى ممهد . ويمكن اجراء هذا التقريب باساليب رياضية او بيانية بهدف التخلص من عدم الانتظام كنتيجة لأخطاء المعاينة من المدرج التكرارى او المضلع التكرارى اللذين يصفان البيانات من العينة . ولذلك يمكن النظر الى المنحنى التكرارى أنه عبارة عن مضلع تكرارى ممهد أى استبدال الخط البيانى المنكسر بمنحنى ممهد

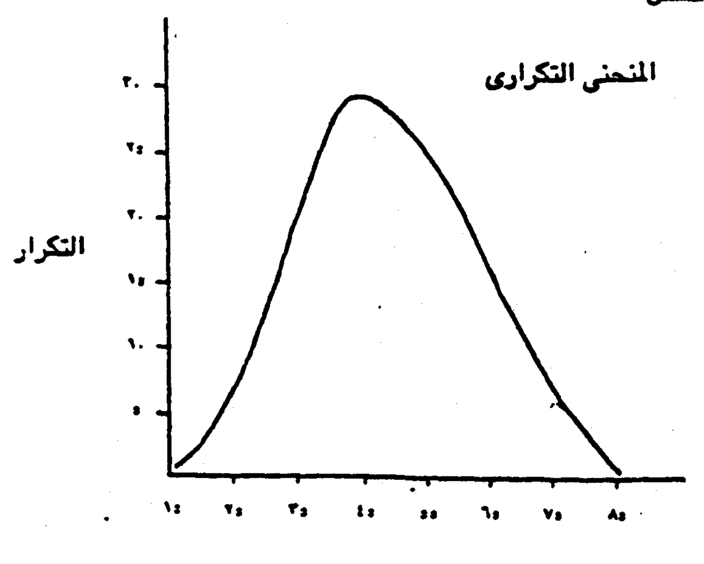
ولذلك عند رسم المنحنى التكرارى نرسم محورين متعامدين ويخصص المحور الأفقى لمراكز الفئات والمحور الرأسى للتكرارات . ثم نرصد النقط على الشكل ونصل بين هذه النقط بمنحنى ممهد ويلاحظ أن المساحة المحصورة بين المحور الأفقى والمنحنى التكرارى تمثل مجموع التكرارات . كما يلاحظ أن المنحنى الممهد يمكن استخدامه فى حالة البيانات المتصلة فقط . كما يمكن استخدام التكرارات النسبية لكل فئة . وفى هذه الحالة يطلق على المنحنى الممهد بمنحنى التكرار النسبى وتصبح المساحة المحصورة بينه وبين المحور الأفقى واحد صحيح . ومن الواضح ان هذه الاجراءات تعتبر صالحة فقط فى حالة تساوى الفئات .

مثال :

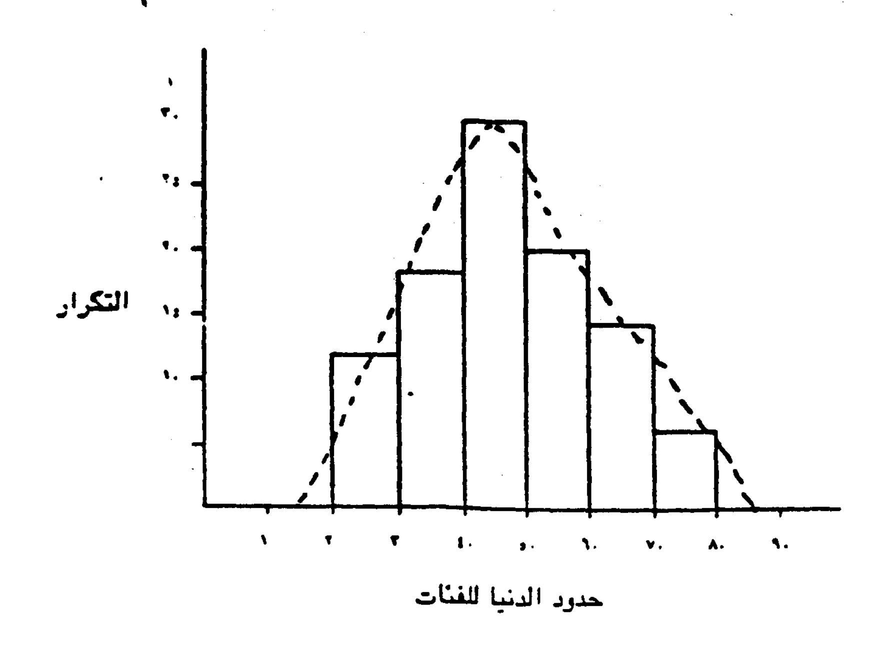
اذاكان لديك التوزيع التكرارى الآتى لمبيعات ١٠٠ شركة بآلاف الجنيهات . والمطلوب رسم المنحنى التكرارى ثم رسم المدرج التكرارى والمنحنى التكرارى فى شكل واحد .

فئـات (المبيعـات)	التكرار (عدد الشركات)
٢٠ -	١٢
٣٠ -	١٨
٤٠ -	٣٠
٥٠ -	٢٠
٦٠ -	١٤
٧٠ -	٦
المجموع	١٠٠

الحـل :



اما رسم المدرج التكرارى والمنحنى التكرارى فى شكل واحد



ويمكن كذلك رسم المدرج التكرارى للتكرارات النسبية وكذلك رسم المنحنى التكرارى للتكرارات النسبية بنفس الأسلوب السابق فيما عدا وضع التكرارات النسبية على المحور الرأسى

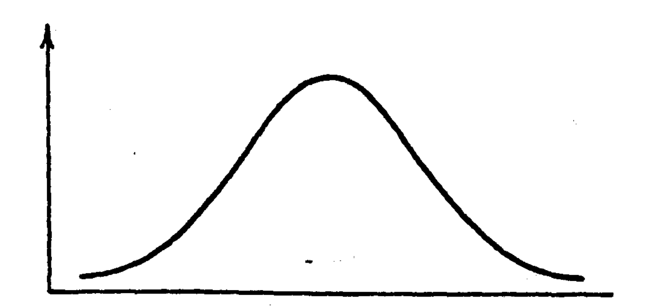
ويلاحظ أننا اوضحنا كيفية انشاء الشكل البيانى للمنحنى التكرارى بنوعيه التكرار المطلق والنسبى وذلك فى حالة وجود توزيعات تكرارية متساوية الفئات إلا إنه فى حالة وجود فئات غير متساوية فانه تتبع نفس القواعد السابقة فيما عدا ضرورة ايجاد كثافة التكرار وتمثيلها على المحور الرأسى بدلا من التكرارات الاصلية .

ويلاحظ عند رسم المنحنى التكرارى بصفة عامة أنه يعتمد على دقة الرسم وقد لا تقع جميع النقط على المنحنى بل يمر المنحنى بينها . ولذلك فان المساحة التى تكون تحت المنحنى قد لا تساوى مساحة المدرج التكرارى أو المضلع التكرارى بالضبط ولكنها تكون قريبة جدا منها .

وهناك انواعا كثيرة من المنحنيات التكرارية أهمها المنحنيات التكرارية وحيدة القيمة . ويمكن تقسيم هذه المنحنيات الى قسمين .

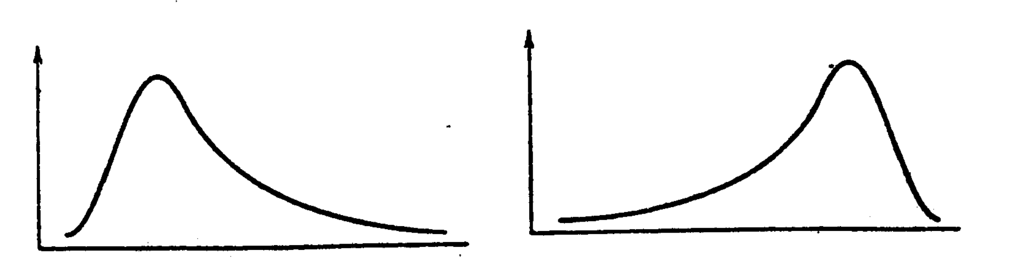
منحنيات متماثلة Symetric Curves وهى المنحنيات التى إذا اسقطنا عمودا من أعلى قمة فى المنحنى على المحور الأفقى فانه يقسم المنحنى الى قسمين متساويين ينطبقان على بعضهما تمام الانطباق وأهم هذه

المنحنيات المنحنى الطبيعى أو المعتدل Normal Distribution ويأخذ الشكل الجرسى كما هو موضح بالشكل الآتى ، كما يأخذ هذا المنحنى خواص رياضية سنذكر فى حينه



منحنيات غير متماثلة (ملتوية) : SKEWED CURVES

وتعرف ايضا بالمنحنيات الملتوية أى التى لا تقسم المنحنى التكرارى الى قسمين متساوين ينطبقان على بعضهما تمام الانطباق ، لذلك لايتساوى انحدار جانبى المنحنى . وقد تكون هذه المنحنيات ملتوية ناحية اليمين أى التوائها موجبا عندما يكون الطرف الأكبر للمنحنى ناحية اليمين أى يصعد بسرعة ويهبط ببطء . ويقال لهذه المنحنيات انها ملتوية ناحية اليسار أى التوائها سالبا عندما يكون الطرف الأكبر للمنحنى ناحية اليسار أى يصعد ببطء ويهبط بسرعة.



رابعا : المنحنى التكرارى المتجمع

CUMULATIVE FREQENCY CURVE

أوضحنا سلفا أن التوزيع التكرارى المتجمع يستخدم فى حالة الرغبة فى معرفة عدد المفردات التى تأخذ قيمة أكبر من قيمة معينة أو أقل من قيمة معينة أو عدد المفردات بين قيمتين معينتين . وقد استخدم اسلوب تجميع التكرارات المطلقة أو التكرارات النسبية إما باستخدام التكرارات المتجمعة الصاعدة أو التكرارات المتجمعة الهابطة . وقد توضح كذلك أنه يجب ان تناظر التكرارات المتجمعة الصاعدة الحدود العليا للفئات وتناظر التكرارات المتجمعة الهابطة الحدود الدنيا للفئات .

ومن الطبيعى يمكن توفيق منحنى تكرارى مطلق أو منحنى تكرارى نسبى طبقا للتكرارات المتجمعة الصاعدة أو طبقا للتكرارات النسبية المتجمعة الصاعدة التى توضح على المحور الرأسى والحدود العليا للفئات على المحور الأفقى ويطلق على المنحنى التكرارى الناتج المنحنى التكرار المطلق المتجمع الصاعد والمنحنى التكرارى النسبى المتجمع الصاعد على التوالى وكذلك الحال يمكن توفيق منحنى تكرارى مطلق أو منحنى تكرارى نسبى صنـ

للتكرارات المتجمعة الهابطة أو طبقا للتكرارات النسبية المتجمعة الهابطة التى توضح على المحور الرأسى والحدود الدنيا للفئات على المحور الأفقى ويطلق على المنحنى التكرارى الناتج المنحنى التكرارى المطلق المتجمع الهابط والمنحنى التكرارى النسبى المتجمع الهابط على التوالى .

ومن المهم القول أن قواعد رسم المنحنى التكرارى المتجمع الصاعد أو الهابط لا تتغير سواء كانت الفئات متساوية أو غير متساوية .

مثال :

يوضح الجدول التالى مبيعات ١٠٠ شركة بآلاف الجنيهات فى عام ١٩٩٠ والمطلوب تصوير هذه البيانات فى شكل منحنى تكرارى متجمع

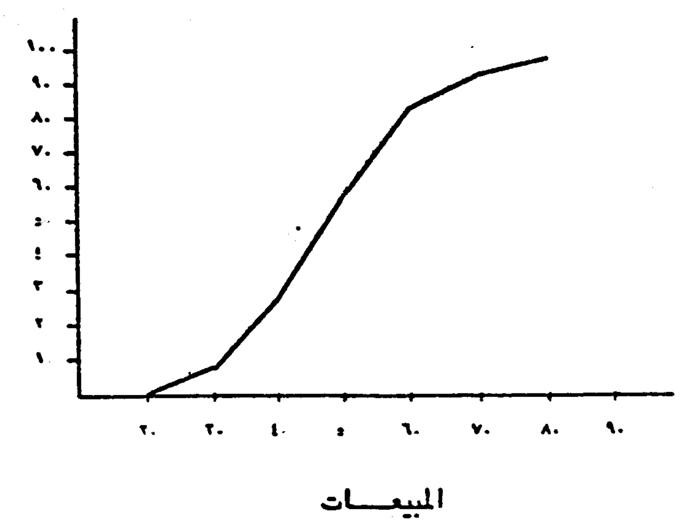
فئات (المبيعات)	التكرار المطلق (عدد الشركات)
٢٠ –	١٠
٣٠ –	٢٠
٤ –	٣٠
٥٠ –	٢٥
٦٠ –	١٠
٧٠ –	٥
المجموع	١

الحـل

المنحنى التكرارى المطلق المتجمع الصاعد

فئات (المبيعات)	التكرار المطلق (عدد الشركات)	الحدود العليا للفئات	التكرار المطلق المتجمع الصاعد
٢٠ –	١٠	أقل من ٣٠	١٠
٣٠ –	٢٠	أقل من ٤٠	٣٠
٤٠ –	٣٠	أقل من ٥٠	٦٠
٥٠ –	٢٥	أقل من ٦٠	٨٥
٦٠ –	١٠	أقل من ٧٠	٩٥
٧٠ –	٥	أقل من ٨٠	١٠٠

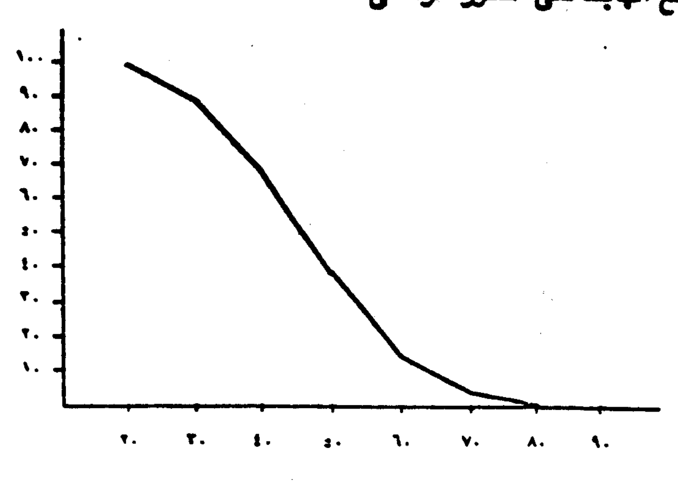
نقوم بعد ذلك بوضع الحدود العليا للفئات على المحور الأفقى والتكرار المطلق المتجمع الصاعد على المحورالرأسى .



المنحنى التكرارى المطلق المتجمع الهابط

فئات (المبيعات)	التكرار المطلق (عدد الشركات)	الحدود الدنيا للفئات	التكرار المطلق المتجمع الهابط
٢٠ -	١٠	٢٠ فاكثر	١٠٠
٣٠ -	٢٠	٣٠ فاكثر	٩٠
٤٠ -	٣٠	٤٠ فاكثر	٧٠
٥٠ -	٢٥	٥٠ فاكثر	٤٠
٦٠ -	١٠	٦٠ فاكثر	١٥
٧٠ -	٥	٧٠ فاكثر	٥

نضع بعد ذلك الحدود الدنيا للفئات على المحور الأفقى والتكرار المطلق المتجمع الهابط على المحور الرأسى .



المبيعـــات

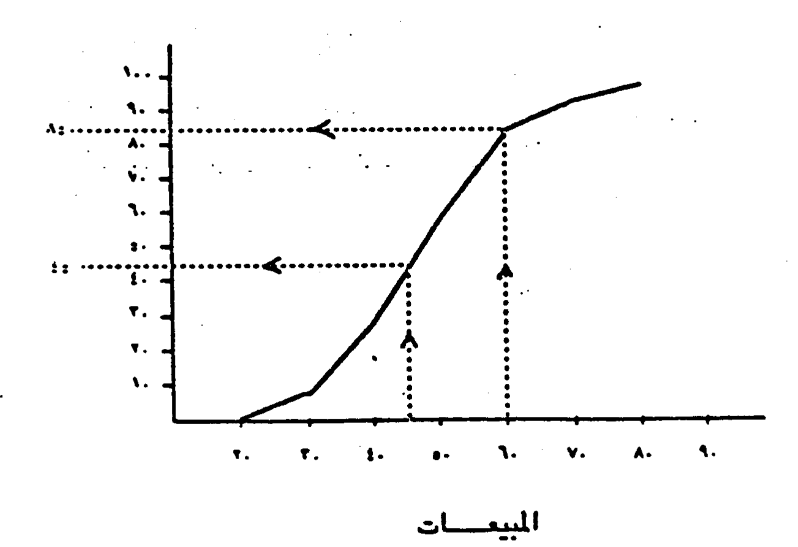
مثـــال :

من الشكل البيانى للمنحنى التكرارى المطلق المتجمع الصاعد أو الهابط المطلوب معرفة عدد الشركات الى تبلغ مبيعاتهاأقل من ٦٠ ألف جنيه . وكذلك عدد الشركات التى تبلغ مبيعاتها أقل من ٤٥ الف جنيه .

الحـــل :

(١) من المنحنى التكرارى المتجمع الصاعد

نرسم عمـودا رأسـيا من النقطة ٦٠ على المحـور الأفـقى حـتى يقـابل المنحنى المتجمع الصاعد ثم نمد خطا مستقيما أفقيا حتى يتقابل مع المحور الرأسى .

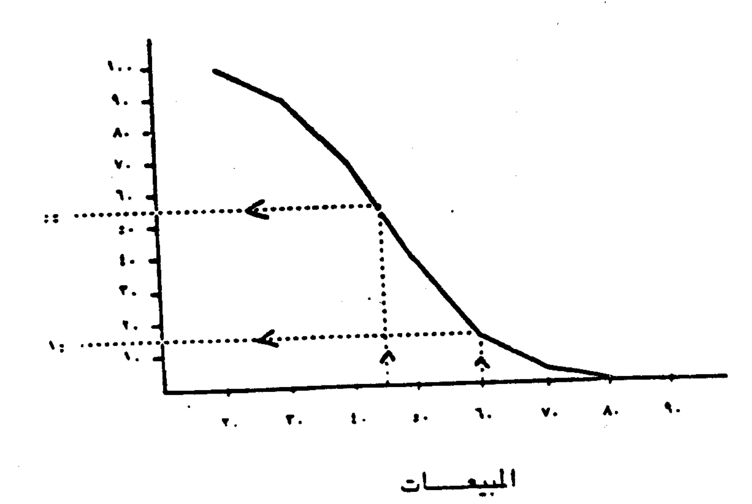


يتضح من الرسم السابق ان عدد الشركات التى تبلغ مبيعاتها اقل من ٦٠ ألف جنيه عدد ٨٥ شركة .

كما نرفع عمودا رأسيا من النقطة ٤٥ على المحور الأفقى حتى يقابل المنحنى المتجمع الصاعد ثم نمد خطا مستقيما أفقيا حتى يتقابل مع المحور الرأسى . ويتضح من الرسم السابق ان عدد الشركات التى تبلغ مبيعاتها اقل من ٤٥ الف جنيه عدد ٤٥ شركة وهكذا .

(٢) من المنحنى التكرارى المتجمع الهابط

نرسم عمودا رأسيا من النقطة ٦٠ على المحور الأفقى حتى يقابل المنحنى المتجمع الهابط ثم نمد خطا مستقيما افقيا حتى يتقابل مع المحور الرأسى ونفس الحال نرفع عمودا رأسيا من النقطة ٤٥ على المحور الافقى حتى يقابل المنحنى المتجمع الهابط ثم نمد خطا مستقيما أيضا حتى يتقابل مع المحور الرأسى .



يتضح من الرسم السابق أن عدد الشركات التى تبلغ مبيعاتها ٦٠ ألف جنيه فاكثر عدد ١٥ شركة . ومن ثم فان عدد الشركات التى تبلغ مبيعاتها أقل من ٦٠ ألف جنيه = ١٠٠ - ١٥ = ٨٥ شركة .

وكذلك يتضح من الرسم السابق ان عدد الشركات التى تبلغ مبيعاتها ٤٥ ألف جنيه فاكثر عدد ٥٥ شركة . ومن ثم فان عدد الشركات التى تبلغ مبيعاتها أقل من ٤٥ الف جنيه = ١٠٠ - ٥٥= ٤٥ شركة .

مثـــال :

افترض ان لديك التوزيع التكرارى السابق اوجد من المنحنى التكرارى المطلق المتجمع الصاعد ومن المنحنى التكرارى المتجمع الهابط عدد الشركات التى تبلغ مبيعاتها بين ٤٥ الف جنيه وستون الف جنيه .

الحـــل :

(١) من المنحنى الكرارى المطلق المتجمع الصاعد

حيث انه سبق ايجاد عدد الشركات التى تبلغ مبيعاتها أقل من ٦٠ الف جنيه وهى ٨٥ شركة .

وحيث انه سبق ايجاد عدد الشركات التى تبلغ مبيعاتها أقل من ٤٥ الف جنيه وهى ٤٥ شركة .

ولذلك فان عدد الشركات التى تبلغ مبيعاتها من ٤٥ ألف جنيه و ٦٠ الف جنيه

= ٨٥ - ٤٥ = ٤٠ شركة

(٢) من المنحنى التكرارى المتجمع الهابط

حيث أنه سبق ايجاد عدد الشركات التى تبلغ مبيعاتها ٤٥ ألف جنيه فاكثر وهى ٥٥ شركة .

وحيث أنه سبق ايجاد عدد الشركات التى تبلغ مبيعاتها ٦٠ الف جنيه فاكثر وهى ١٥ شركة .

ولذلك فان عدد الشركات التى تبلغ مبيعاتها بين ٤٥ الف جنيه و ٦٠ الف جنيه = ٥٥ - ١٥ = ٤٠ شركة .

العرض البياني للبيانات الخام (غير المبوبة):

تختلف وتتعدد العرض البياني للبيانات الخام (غير المبوبة) ولكنها تنحصر في طريقتين أساسيتين:

الطريقة الأولى: وهي طريقة التمثيل البياني عن طريق استخدام الأشكال البيانية.

الطريقة الثانية: وهي طريقة التمثيل البياني عن طريق استخدام الرسوم التصويرية.

وفيما يلي دراسة تفصيلية لكل طريقة على حدة.

أولاً: طريقة العرض البياني بالأشكال البيانية

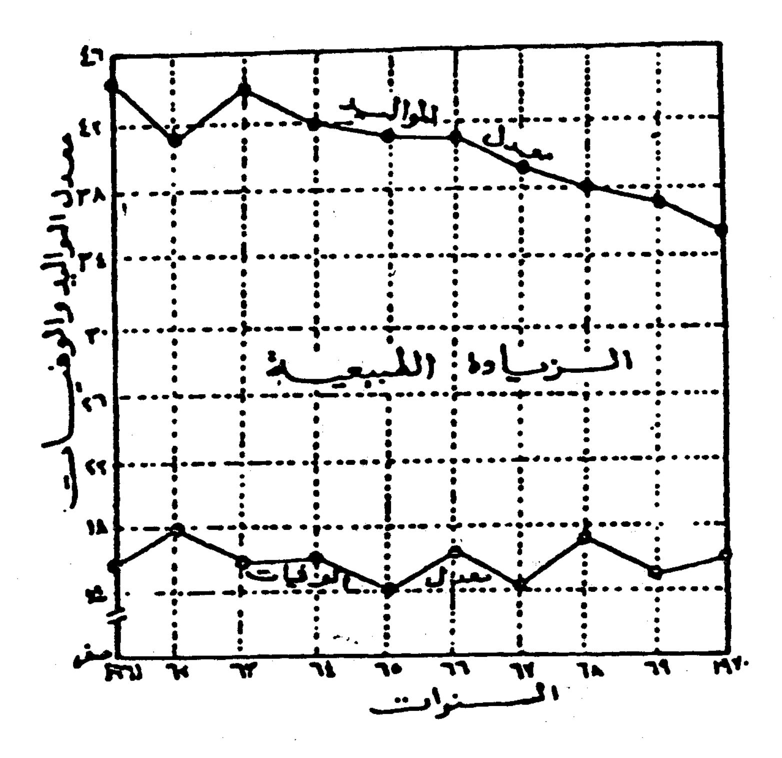
يعتمد هذا الأسلوب في العرض البياني على تمثيل البيانات المتاحة للظاهرات موضع الدراسة في شكل رسوم بيانية انفرادية (أي ليس لها علاقة بالمكان أو الحيز). مثل الخط البياني الذي يمثل الاتجاه العام للظاهرة، الأعمدة البيانية، الرسوم المساحية، الرسوم، الحجمية، الرسوم الثانية أهرامات السكان.

١ ـ الخطوط البيانية Line - graphs

تستخدم الخطوط البيانية في تمثيل التغير من فترة إلى أخرى للظاهرة الواحدة أو لبيان علاقة متغيرين وغالباً ما يكون أحد هذين المتغيرين هو الزمن الذي يعتبر متغيراً مستقلاً. ويبين التغير أو العلاقة بمنحنى، ويظهر ضعف أو شدة التغير في التغير من فترة إلى أخرى ويكون ذلك ما يعرف بالسلسلة الزمنية أو يوضح اتجاه العلاقة بين متغيرين أو أكثر. ولقد جرت العادة عند التمثيل البياني للسلاسل الزمنية أي يكون المحور الأفقي (السيني) ممثلاً للمتغير المستقل (الزمن) والمحور الرأسي

(الصادي التابع) (الظاهرة المدورسة) وفي ضوء البيانات المتاحة يختار مقياس رسم ملائم لأبعاد المسطح المخصص لعملية التمثيل البياني حتى يمكن توقيع كل قيم المتغير التابع على الرسم، فيقسم المحور الرأسي إلى وحدات حسابية بادئين بالصفر ومنتهين بقيمة أكبر من أكبر قيمة تمثل المتغير التابع ويختار كذلك مقياس مناسب للمتغير المستقل (الزمن) على المحور الأفقي. ويجب مراعاة عدم وجود تفاوت كبير في الأبعاد القياسية للمتغيرين حتى لا يؤدي ذلك إلى عدم الدقة. ويتم رسم المنحنى من خلال توقيع جميع القيم على الرسم في شكل نقط تحدد كل منها بأحداثين (أي على حسب بعدي كل نقطة عن المحورين) ثم توصل مواقع القيم فتعطي لنا الشكل المطلوب (الخط البياني) كما في الشكل رقم (٣ ـ ١). ويجب ملاحظة أنه عدم رسم الخطوط لظاهرة متغيرة بانتظام أو تدريجياً أن يكون الخط البياني منحنياً مثل الخط البياني الذي وضع المتوسط الشهري لدرجة الحرارة خلال شهور على مدينة الإسكندرية. وفي بعض الحالات قد لا يبدأ المقياس على المحور الرأسي بالصفر ولكن يبدأ بقيمة أكبر تبعاً لأن البيانات المراد تمثيلها بيانياً بخط بياني تبدأ بقيمة بعيدة عن الصفر ولكن تقترب من بعضها بمدى صغير فإذا ما أخذنا وحدات قياسية تناسب أصغر وأكبر رقم بادئين بالصفر فإن ذلك سيؤدي إلى وجود فراغ كبير غير مستخدم يقع بين الصفر وأصغر رقم موجود مما يترتب عليه أن يجعل الخط البياني محصوراً في أعلى جزء من الرسم وهذا شيء غير مستحب أو مرغوب فيه. وللتغلب على ذلك فإننا نحاول أن نضغط المسافة على المحور الرأسي بين الصفر وأصغر فيه بأن نكسر المحور الرأسي بخطين مائلين بعد نقطة الصفر على المحور الرأسي كما يظهر في الشكل رقم (٣ ـ ١).

وفي حالة إذا كنا بصدد تمثيل سلسلة زمنية لبيانات تتفاوت القيم فيها تفاوتاً كبيراً أو في شكل معدلات مثل معدل النمو أو التغير السكاني من سنة لأخرى، معدل التغير في الاستهلاك معدلات التغير في الدخل القومي، أو نسب تطور الدعم الحكومي للسلع والخدمات، أو نسب النقص والزيادة في أي ظاهرة فإنه يجب أن يقسم المحور الرأسي إلى وحدات لوغاريتمية بدلاً من الوحدات الحسابية.



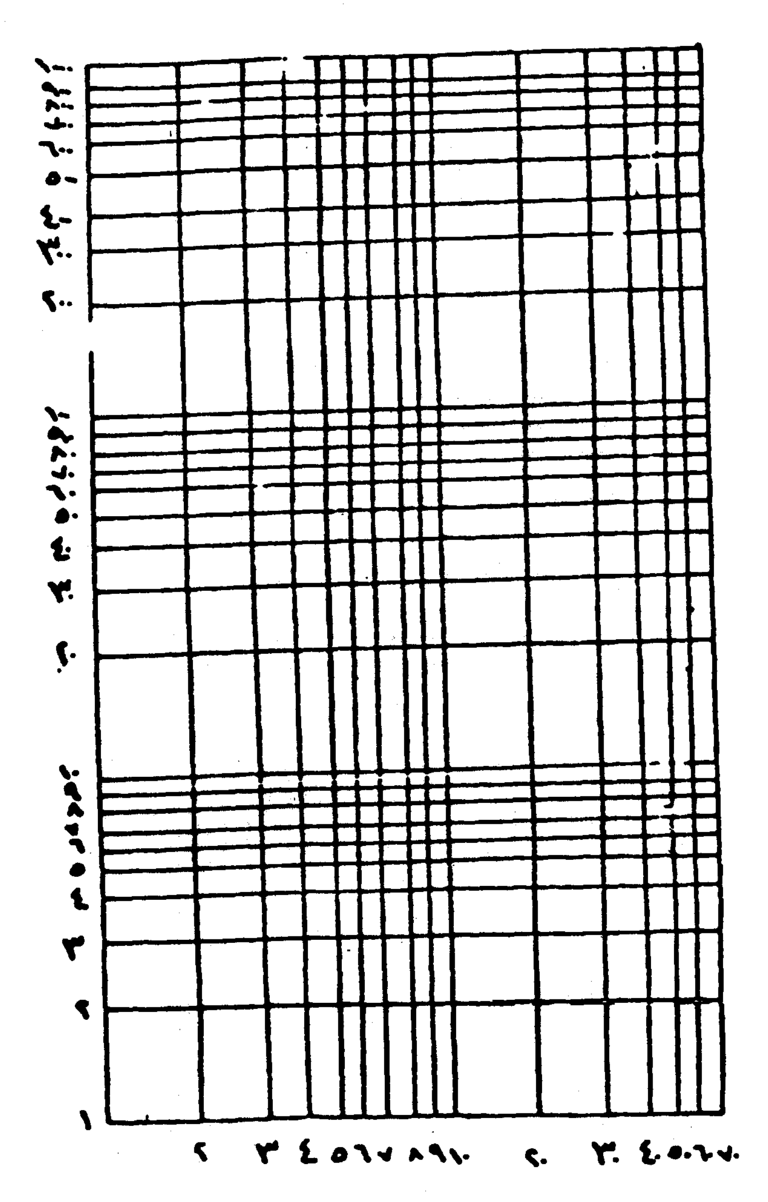
شكل رقم (٣ ـ ١): الخطوط البيانية الحسابية

وتقوم فكرة التقسيم اللوغاريتمي على أخذ لوغاريتم الأعداد من ١ إلى ١٠ وجعلها أساساً لوحدة التقسيم اللوغاريتمي والتي تضرب في كل مرة في طول الدورة اللوغاريتمية المأخوذة طبقاً لطول المسافة الرأسية والأفقية المراد تمثيل الظاهرة عليها، وهي في هذه الحالة تمثل دورة لوغاريتمية واحدة. وبعد ذلك يمكن أيضاً أخذ دورة لوغاريتمية ثانية تبدأ بالرقم ١٠ حتى الرقم ١٠٠ وتأخذ نفس قياسات الدورة الأولى (١ ـ ١٠) كما يمكن أخذ دورة لوغاريتمية ثالثة تبدأ بالرقم

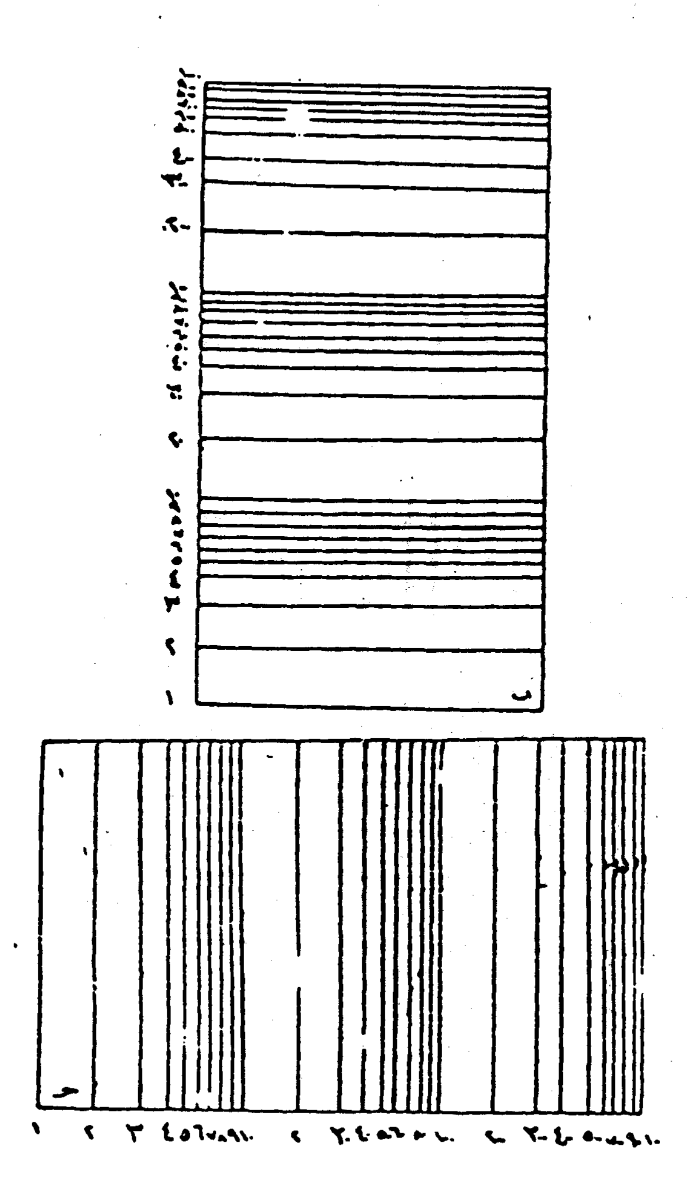
١٠٠ وتنتهي بالرقم ١٠٠٠ ولها نفس قياسات الدورة الأولى أيضاً وتمثل مئات أضعاف الوحدة الحسابية الواحدة. فإذا فرض أنه كانت لدينا بيانات أصغر رقم فيها هو ٢٠ وأكبر رقم ٣٠٠٠ فإنه يكفي لتمثيل هذه البيانات على رسم بياني لوغاريتمي مكون من ثلاث دورات، تمثل الدورة الأقسام من ١٠ حتى ١٠٠ والثانية من ١٠٠ حتى ١٠٠٠ والثالثة من ١٠٠٠ حتى ١٠٠٠٠ ويؤخذ طول الدورة الواحدة مساوياً لخمسة سنتيمترات وقد جرت العادة على أن يبدأ التقسيم اللوغاريتمي بالرقم (١) مع قسمته على أو ضربه في الرقم ١٠ أو مضاعفاته حتى يمكن البدء بالأرقام ١ر٠ أو ١ر٠ أو ١ر٠ كما يمكن البدء بالرقم ١٠ أو ١٠٠ أو ١٠٠٠ وهكذا.

وكما في رسم الخطوط البيانية الحسابية وعلى حسب البيانات المتاحة يمكن أن يقسم المحور الرأسي فقط تقسيماً لوغاريتميا لتوقع على أساسه معدلات التغير في ظل الفترة الزمنية التي يقسم على أساسها المحور الأفقي (يسمى ذلك بالتقسيم نصف لوغاريتمي). كما قد يقسم كل من المحورين الأفقي والرأسي تقسيماً لوغاريتميا (يطلق عليه اسم التقسيم اللوغاريتمي المزدوج) ويناسب ذلك البيانات التي تتكون من معدلات تغير أو نسب مئوية لمتغيرين مستقيم. وتبعاً لأهمية استخدام التقسيم اللوغاريتمي في التمثيل البياني فإنه يوجد حالياً ورق رسم بياني خاص مقسم تقسيمة الوغاريتميا أما على المحور الأفقي أو الرأسي أو مزدوجاً كما توضحه الأشكال الآتية (شكل زقم ٣ ـ ٢ أ، ب، جـ).

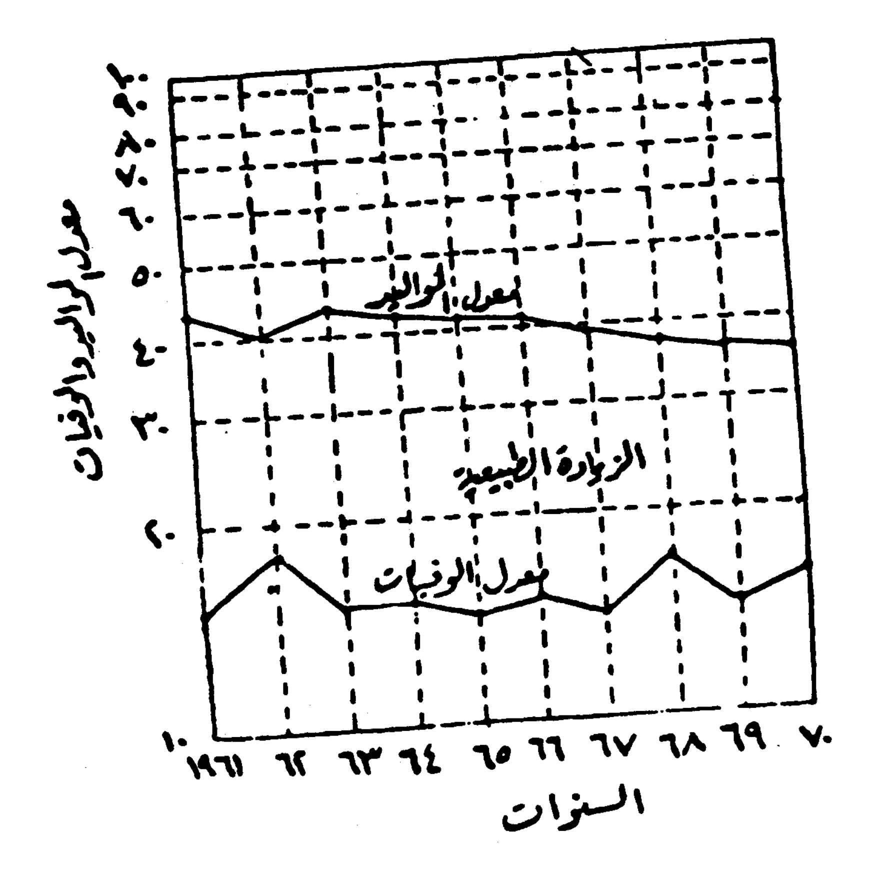
والشكلان رقم (٣ ـ ١)، (٣ ـ ٣) يوضحان مقارنة بين الخطوط البيانية الحسابية واللوغاريتمية لمعدلات المواليد والوفيات في جمهورية مصر في الفترة من ١٩٦١ ـ ١٩٧٠، ومنها يتضح أن الخطوط البيانية اللوغاريتمية التي رسمت على رسم بياني نصف لوغاريتيم لا تظهر حدة التغير في كل من معدلات المواليد والوفيات التي أظهرتها الخطوط البيانية الحسابية وانعكس ذلك أيضاً على ازيادة الطبيعية للسكان المحسوبة من الرسم في كل من الشكلين.



شكل رقم (٣ ـ٢ أ) ورقة بيانية لوغاريتمية
(مزدوجة)



شكل رقم (٣ ـ ٢ ب، جـ) ورق رسم بياني لوغاريتمي
ب أنفي، جـ ـ رأسي

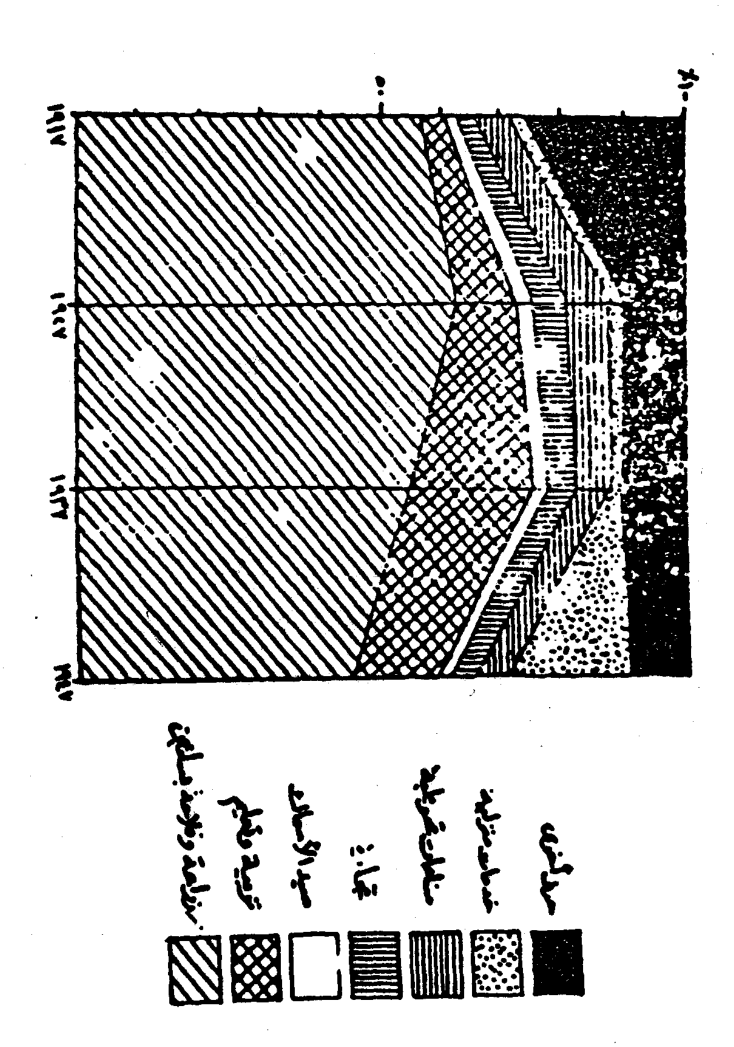


شكل رقم (٣ ـ ٣) الخطوط البيانية اللوغاريثمية

وهناك نوع آخر من الخطوط البيانية يعرف باسم المنحنيات المجمعة Compound Curve Graph وهي عبارة عن خطوط بيانية تمثل التغير في مجموع الظاهرة. وظاهرة (أو ظاهرات) أخرى، بحيث يمثل التغير في أجزاء تنقسم إليها هذه الظاهرة ثم تظلل المساحات المحصورة بين ذه الخطوط البيانية (شكل: رقم ٣ ـ ٤). ويمكن رسم هذا النوع من الرسوم على أساس النسب المئوية وهي بطبيعة الحال تكون الأنسب والأحسن ويتم ذلك بتقسيم الظاهرة إلى أجزائها المختلفة بشرط أن تكون بنفس الترتيب لكل فترة زمنية ثم نصل بين نقط التقسيم بخطوط وتظلل المساحات المحصورة بين هذه الخطوط وبذلك يمكن معرفة عما إذا كانت نسبة أي قسم من الظاهرات قد هبطت أو زادت في نفس الوقت بالنسبة إلى باقي التقسيمات الفرعية للظاهرة. ويطلق على هذا النوع بصفة عامة اسم الرسوم البيانية المجمعة Compound line or Band graph.

٢ ـ الأعمدة البيانية Bar Graph

تعتبر طريقة الأعمدة البيانية من أبسط طرق التمثيل البياني التي تستخدم للمقارنة بين الكميات لظاهرة واحدة أو عدة ظواهر، وعادة ما تسمى رسومها البيانية باسم Columnar diagrms وتتألف هذه الرسوم من أعمدة ذات عرض متساوى وطول تناسب مع الكميات التي تمثلها حسب مقياس الرسم المختار. ويمكن رسم هذه الأعمدة الرأسية لها رأسياً وأفقياً في أشكال بيانية قائمة بذاتها، وتعتبر الأعمدة الأفقية أفضل من حيث سهولة قراءتها، أما الأعمدة الرأسية فلها ميزة أخرى وهي سهولة المقارنة. وقد تكون هذه الأعمدة بسيطة حينما يرسم كل عمود منها لكي يوضح المجموع الكلي فقط، أو قد تكون مركبة Compound حينما يقسم كل عمود لكي يبين التقسيمات الفرعية إلى جانب المجموع الكمي.

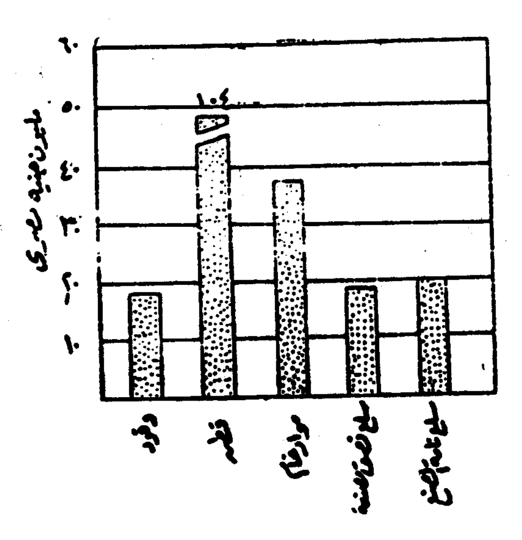


شكل رقم (٣ ـ ٤) الخطوط البيانية المجمعة

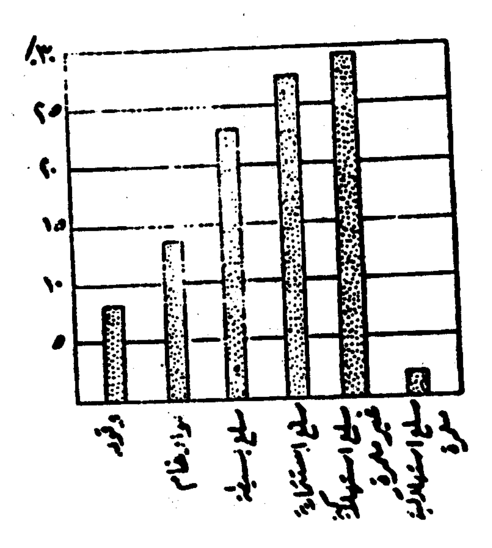
وتعتمد طريقة الأعمدة البيانية البسيطة على تمثيل البيانات الوصفية وفي إظهار كميات الزيادة والنقص في بعض الظواهر مثل معدلات المواليد والوفيات أو تمثيل التطور في أي سلسلة زمنية خاصة بالإنتاج أو الاستثمار أو حجم المشتريات أو الدخل أو عدد السكان خلال فترات زمنية معينة. وفي هذه الحالة نرسم محورين أحدهما محور رأسي يقسم إلى أقسام متساوية حسب الفترات الزمنية أو الصفات المميزة للظاهرة كالحالات التعليمية أو الاجتماعية أو فئات السن. . . إلخ. ومما هو جدير بالذكر أنه عند أخذ المسافات الممثلة لقواعد الأعمدة على المحور الأفقي يجب أن تكون متساوية وعلى أبعاد متساوية أيضاً، ذلك بطريقة تلائم المساحة من لوحة الرسم المخصصة للتمثيل البياني وعدد الأعمدة المراد رسمها. وفي حالة الفترات الزمنية غير المتنظمة فإن المسافات بين كل عمود وآخر يجب أن يتناسب مع الأبعاد الزمنية للفترة المراد تمثيلها بيانياً، كما يجب في كل الحالات أن يبدأ المقياس على المحور الرأسي من الصفر وينتهي برقم أعلى من أكبر قيمة من قيم الظاهرة موضع البحث. إلا أنه في كثير من الحالات نجد بين قيم الظاهرة المراد تمثيلها بطريقة الأعمدة البسيطة قيمة أو قيمتين متطرفتين أو شاذين تفوق بقية قيم الظاهرة مما يؤدي إلى وجود تفاوت كبير لهذه القيم. وبالتالي يؤدي ذلك إلى اختلاف كبير في طول الأعمدة، بل أنه في بعض الأحيان يصبح من الصعب تمثيل القيم بأخذ مقياس رسم على المحور الرأسي فيلائم هذه القيم المتفاوتة. فمثلاً إذا كانت لدينا كمية أكبر مائة مرة من كمية أخرى، فإنها تتطلب رسم عمود أطول مائة مرة من عمود الكمية الأصغر، وهذا يضطرنا إلى أن نرسم الكميات الصغيرة بأعمدة صغيرة جداً. وأما أن نرسم أعمدة قد يضطرنا طولها الكبير جداً إلى تقطيعها توضع بجوار بعضها البعض. ولو أن كل هذا التحايل لا ينقل الصورة الصحيحة لتمثيل هذه الكميات، وما لذلك من تقليل من أهمية هذا الأسلوب في التمثيل البياني. وللتغلب على هذه المشكلة يستحسن قطع المحور الرأسي الموجود عليه لقياس الكميات وجعل الجزء الأسفل منه يبدأ من الصفر حتى قيمة أعلى من الكمية الصغيرة، أما الجزء الأعلى فيبدأ من رقم أقل من الكمية الكبيرة وينتهي برقم أعلى

منها مع ثبات طول المقياس في الجزئين. وفي بعض الأحيان تكسر الأعمدة التي تمثل قيما متطرفة ويكون ذلك بالتخلص من الارتفاعات التي تعلو الارتفاعات العادية للقيم الأخرى وقد يكون كسر الأعمدة رأسياً عن طريق وضع خطين متوازيين مائلين عند نهاية الارتفاع المراد تحديده والذي يناسب الشكل رليدل على أن للعمود بقية ولكن مساحة ورقة الرسم لا تسمح بإظهارها، ولكن يجب أن نكتب أعلى هذا العمود بالذات الكمية الحقيقية التي يمثلها (شكل رقم ٣ ـ ٥) وعلى الرغم من ذلك فإن هذه الطريقة لا يمكن الاستفادة بها في حالة المقارنة لأنها لا تظهر الفرق بين الكميات كحقيقتها. ولكن هناك نوع آخر من الأعمدة البيانية يصلح في إظهار الأهمية النسبية لمكونات الظاهرة يسمى بالأعمدة النسبية Proportional Bars والتي ترسم لتمثيل إعداد السكان أو الإنتاج المعدني أو حركة الصادرات والواردات في الموانىء وتتميز طريقة الأعمدة النسبية بسهولة رسمها من ناحية التصميم، وكذلك بسهولة القرار من الناحية المرئية. وتتلخص طريقة رسم الأعمدة النسبية في أن نبدأ أولاً باستخراج النسبة المئوية للكميات التي نريد تمثيلها بالنسبة للمجموع الكلي للكميات مثل نسبة وزن المجموعات الرئيسية للواردات المصرية في سنة ١٩٦٢، ويخصص المحور الأفقي لتعيين المجموعات المختلفة للواردات. أما المحور الرأسي فيخصص للأوزان المناظرة لكل مجموعة ويقسم إلى أقسام متساوية تبين النسبة المئوية للأوزان مبتدئين بالصفر ومنتهي بنسبة أعلى م أعلى النسب المراد تمثيلها كما في الشكل رقم (٣ ـ ٦) ويحسن عند رسم هذا النوع من الأعمدة أو نختار له نوع التظليل المصمت (كاللون الأسود المصمت) أو نستخدم نمط التظليل النقطي وذلك لأن استخدام نمط الخطوط المائلة في تظليل فراغ الأعمدة يتضمن نوعاً من خداع البصر.

أما طريقة الأعمدة البيانية المركبة Compound - Bar (شكل رقم ٣ ـ ٧) فهي عبارة عن أعمدة ذات عرض متساوية ومقسمة إلى أقسام داخلية تمثل في مجموعها المجموع الكمي للظاهرة وفي هذه الحالة فإنه يمكن مقارنة كميات هذه الظاهرة من ناحية الكميات المطلقة ـ كما أنه يمكن مقارنة كميات هذه الظاهرة من الناحية

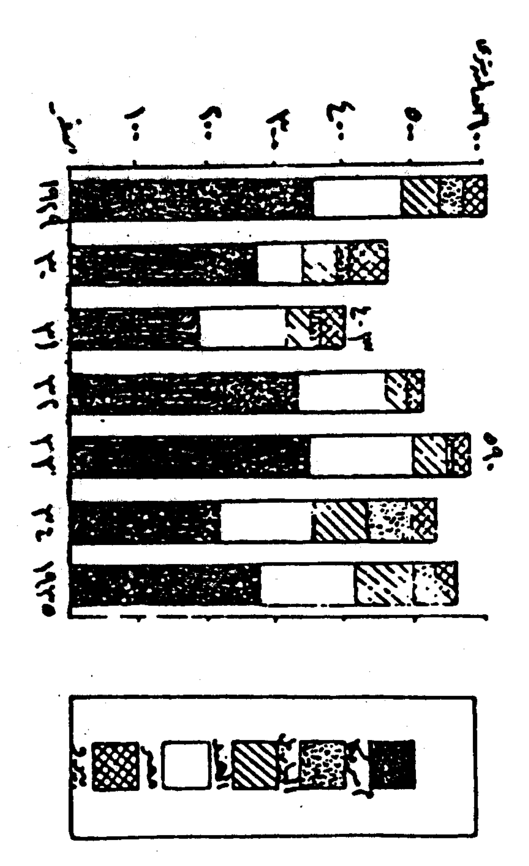


شكل رقم (٣ ـ ٥) الأعمدة البيانية البسيطة

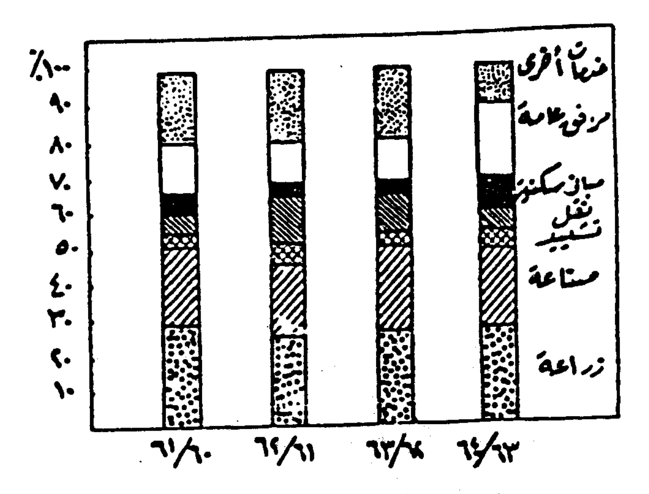


شكل رقم (٣ ـ ٦) الأعمدة البيانية النسبية

النسبية وذلك بتحويل كمية كل قسم فرعي منها إلى نسبة مئوية. وتسمى الأعمدة البيانية في هذه الحالة باسم الأعمدة المركبة النسبية. وفي هذه الحالة لا يمكن مقارنة كل عمود (مستطيل) بآخر ولكن يمكن بمفارنة الجزئيات (التفاصيل) ومن كل عمود بالجزء الذي يناظره في العمود الآخر، وذلك بمعرفة الفرق بين نسبتيهما بالنسبة للمجموع الكمي. ويجب في هذه الحالة أن يصاحب الرسم البياني للأعمدة المركبة النسبية رسم بياني آخر تكون أرقامه مطلقة حتى يمكن معرفة التغير بين المجموع الكمي لكل ظاهرة وأخرى (شكل رقم ٣ ـ ٨).



شكل رقم (٣ ـ ٧) الأعمدة البيانية المركبة المطلقة



شكل رقم (٣ ـ ٨) الأعمدة البيانية المركبة النسبية

٣ ـ الرسوم البيانية المساحية Areal Graphs

تستخدم الرسوم البيانية المساحية لتمثيل البيانات التي يوجد فيها تفاوت كبير بين أرقامها والتي لا يمكن تمثيلها بالخطوط أو الأعمدة البيانية وذلك لأنها تدخل في حسابها البعد الثاني (المساحة). وأوضح أنواع الرسوم البيانية المساحية هي الدائرة والمربع، ولكن لما كانت الدائرة أسهل كثيراً في رسمها فهي أكثر شيوعاً واستخداماً من المربع حيث أنها تشغل حيزاً أقل، كما أنه يمكن تقسيم الدائرة إلى أقسام متعددة حسب ما تجنيه الظاهرة من تفاصيل.

وتعتمد رسم الدوائر البيانية، والتي تستخدم لبيان ومقارنة ظاهرتين أو أكثر أو مقارنة ظاهرة واحدة بنوعياتها خلال فترات زمنية متفاوتة، على إظهار التفاوت بين المجموع الكمي لقيم الظاهرة أو من ظاهرة إلى أخرى وهذا لا يتأتى إلا إذا

قمنا برسم دوائر ذات أنطار متساوية لأن ذلك لا يحدث فقط إلا إذا تساو المجموع الكمي لكل ظاهرة ويمكن أن نستفيد من استخدام هذه الدوائر في حالتين أساسيتين عندما يكون المجموع الكلي كبيراً نسبياً ولكنه يمثل في مساحة محدود جداً ـ كما في حالة تمثيل عدد سكان المدن أو تمثيل ... المصانع، أو عندما نريد تمثيل الكميات الكلية في منطقة أو أقليم أو دولة كما في حالة تمثيل إنتاج البترول في البلاد العربية مثلاً.

ونظراً لأن مساحة الدائرة تتناسب مع مربع نصف قطرها (مساحة الدائرة = ط نق^{2}) فإنه يجب عند رسم الدوائر البيانية أخذ الجذر التربيعي للقيم الكلية التي تمثلها أما إذا كان التمثيل البياني لظاهرة واحدة فقط فعلينا أن نختار الطول المناسب والذي يتلائم مع مساحة ورقة الرسم المراد تمثيل الظاهرة عليها.

ولتمثيل الإحصائية الآتية بطريقة الدوائر البيانية نجرى الآتي: ـ

إنتاج مناطق الصيد بجمهورية مصر في
الفترة من ١٩٦١ ـ ١٩٦٤ ـ (بالألف طن)

السنوات / الإنتاج	٦٢/٦١	٦٣/٦٢	٦٤/٦٣
إنتاج البحار	٦٥	٦٥	٧٠
إنتاج البحيرات	٤٧	٥٤	٥٦
إنتاج النيل	١٥	١٦	١٨
المجموع	١٢٧	١٣٥	١٤٤

أ ـ نجمع الإنتاج في كل سنة حتى نحصل على المجموع الكلي (السنوي) للإنتاج في كل فترة زمنية.

ب ـ نستخرج الجذر التربيعي لمجموع الإنتاج في كل سنة على حدة ويكون الناتج ممثلاً لطول نصف القطر (نق) الذي نريد أن نعرفه لكي نرسم الدوائر التي تمثل الإنتاج، والجذور التربيعية للمثال هي ٣ر١١، ٦ر١١، ١٢.

جـ ـ تختار قيمة قياسية أساسية سواء بالسنتيمتر أو المليتر، يمكن على أساسها أن نحول أعداد الجذور التربيعية الناتجة لدينا إلى أطوال متناسبة تمثل مباشرة أنصاف أقطار الدوائر. وفي العادة تعطي هذه القيمة الأساسية لأصغر جذر تربيعي.

د ـ لمعرفة أنصاف أقطار الدوائر هناك عدة طرق تؤدي إلى نتيجة واحدة ولكنها تختلف في العمليات الحسابية. وسنختار من هذه الطرق طريقتين مألوفتين هما: طريقة التناسب الحسابي «طريقة المقص» ويمكن أن نطبقها على المثال السابق فمثلاً إذا اخترنا الطول ١٦ مليمتر كقيمة أساسية للجذر التربيعي ٣ر١١ فإن:

٣ر١١ = ١٦ مليمتر (نصف قطر الدائرة الأولى)
٦ر١١ = س مليمتر (نصف قطر الدائرة الثانية)

$$س = \frac{١٦ \times ٦ر١١}{٣ر١١} = ٤ر١٦ \text{ مليمتر}$$

٣ر١١ = ١٦ مليمتر (نصف قطر الدائرة الأولى)
١٢ = س مليمتر (نصف قطر الدائرة الثالثة)

$$س = \frac{١٢ \times ١٦}{٣ر١١} = ١٧ \text{ مليمتراً}$$

والقيمة الأساسية التي اخترناها يعتمد اختيارها على مساحة لوحة الرسم ويجب عند اختيارها أن تتوافق مساحة الدوائر مع أبعاد مسطح الرسم بحيث لا تظهر أصغر دائرة صغيرة جداً. وأكبر دائرة كبيرة جداً بالنسبة لمساحة لوحة الرسم.

أما الطريقة الأخرى فهي طريقة سهلة ولا تتطلب كثيراً من الحساب وتتلخص في أن نقسم الجذور التربيعية على العدد ١٠ أو قوى هذا العدد (١٠٠، ١٠٠٠، ١٠٫٠٠٠... الخ) وذلك طبعاً على حساب المدى الذي توجد عليه الجذور التربيعية. ففي مثالنا السابق يمكن أن نقسم الجذور التربيعية كلها على العدد ١٠ ويكون تمييز الناتج بالسنتيمتر وعلى هذا الأساس نجد أن:

نصف قطر الدائرة الأولى = ١١٫٣ ÷ ١٠ = ١٫١٣ سنتيمتر

نصف قطر الدائرة الثانية = ١١٫٦ ÷ ١٠ = ١٫١٦ سنتيمتر

نصف قطر الدائرة الثالثة = ١٢ ÷ ١٠ = ١٫٢ سنتيمتر

وكما هو واضح فإن الأطوال التي نتجت بهذه الطريقة هي أطوال صغيرة وبالتالي ستكون مساحات دوائرها صغيرة، وفي مثل هذه الحالة يجب أن نكبر الأطوال الناتجة، وذلك بضربها كلها في أي رقم نختاره، بحيث تظهر الدوائر بعد رسمها ملائمة لأبعاد لوحة الرسم. فإذا كان هذا الرقم الذي اخترناه هو ١٫٥ مثلاً فسوف يصبح:

طول نصف قطر الدائرة الأولى = ١٫١٣ × ١٫٥ = ١٫٦٩٥ سم

(١٫٧ سم تقريباً)

طول نصف قطر الدائرة الثانية = ١٫١٦ × ١٫٥ = ١٫٧٤٠ سم

(١٫٧٥ سم تقريباً)

طول نصف قطر الدائرة الثالثة = ١٫٢ × ١٫٥ = ١٫٨٠ سم.

وأنصاف الأقطار السابقة هي التي رسمت على أساسها الدوائر في الشكل رقم (٣ ـ ٩). وسواء استخدمنا أي من الطريقتين السابقتين لمعرفة أطوال نصف قطر الدوائر فيجب أن لا نكتب عليها أية أعداد للكميات الحقيقية التي تمثلها

الدوائر. . . وسيتبع ذلك أن يرسم في أحد أركان لوحة الرسم منتاح قياس الدوائر بنفس طريقة رسم الدوائر السابق شرحها ومنه يمكن أن نقيس قطر أي دائرة مرسومة وليس من الضروري أن يمثل دوائر المقياس نفسها وإنما يمثل مقياساً لدوائر كمياتها ذات أرقام صحيحة دائرية بحيث تكون قريبة من الكميات الحقيقية التي تم تمثيلها بيانياً. وفي المثال السابق فإن هذه الكميات ١٠٠٠، ١٥٠، ٢٠٠ (شكل رقم ٣ ـ ٩).

ويمكن تقسيم الدائرة إلى أقسام داخلية للمقارنة بين أجزاء الظاهرة. وفي هذه الحالة تحول الأرقام المطلقة إلى أرقام نسبية عن طريق قسمة رقم كل جزء على المجموع الكلي للظاهرة وضربة في ١٠٠ ثم ضرب الناتج في ٦ر٣ فنحصل على الزاوية المركزية التي تمثل هذا الجزء وفي المثال السابق تقسم الدائرة الأولى (٦١/٦٢) إلى ثلاثة أقسام بنسبة إنتاج البحار والبحيرات والنيل كما يلي: ـ

نسبة إنتاج البحار = $\frac{٦٥}{١٢٧}$ × ١٠٠ = ٢ر٥١٪

نسبة إنتاج البحيرات = $\frac{٦٥}{١٢٧}$ × ١٠٠ = ٣٧٪

نسبة إنتاج النيل = $\frac{١٥}{١٢٧}$ × ١٠٠ = ٨ر١١٪

وتكون الزاوية المركزية لكل منها هي:
إنتاج البحار = ٢ر٥١ × ٦ر٣ = ١٨٤°
إنتاج البحيرات = ٣٧ × ٦ر٣ = ١٣٣°
إنتاج النيل = ٨ر١١ × ٦ر٣ = ٤٢°

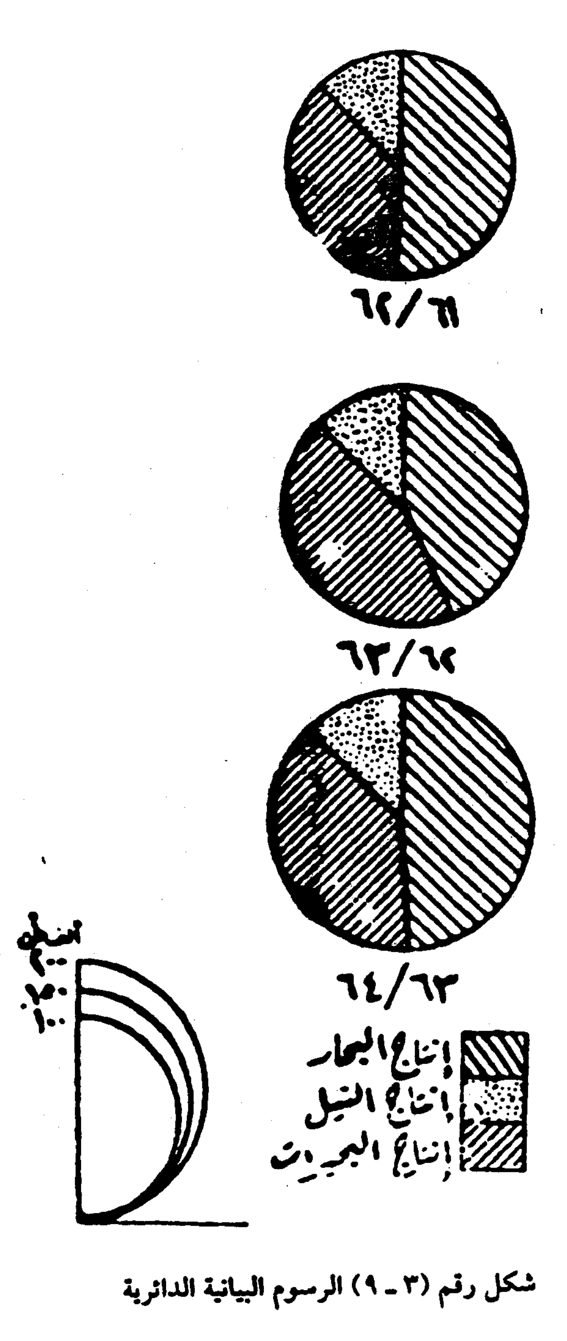
وبالمثل حساب النسبة المئوية والزاوية المركزية للإنتاج المصايد في السنتين الأخيرتين كما في الجدول التالي:

السنوات / إنتاج	٦١/٦٢		٦٢/٦٣		٦٣/٦٤	
	٪	الزاوية المركزية	٪	الزاوية المركزية	٪	الزاوية المركزية
إنتاج البحار	٥١ر٢	١٨٤	٤٨	١٧٢ر٨	٤٨ر٦	١٧٥
إنتاج البحيرات	٣٧	١٣٣	٤٠	١٤٤	٣٨ر٩	١٤٠
إنتاج النيل	١١ر٨	٤٢	١٢	٤٣ر٢	١٢ر٥	٤٥
المجموع	١٠٠	٣٦٠	١٠٠	٣٦٠	١٠٠	٣٦٠

ولتسهيل المقارنة بين أجزاء الظاهرة خلال فترة السنوات الثلاث يجب أن تأخذ الضلع الرأسي المكون للربع الأول من الدائرة (الخط الواصل بين مركز الدائرة وبداية تقسيم الدائرة أو صفر التدريج) كخط أساسي سيبدأ منه قياس الزوايا المركزية بعد تجميعها تصاعدياً. (شكل رقم: ٣ ـ ٩).

وينطبق كل ما ذكرناه في طريقة رسم الدوائر البيانية على طريقة رسم المربعات، فكلاهما صالح لنفس الاستخدام لتمثيل البيانات. وتستخدم المربعات في الحالات التي يراد فيها التنويع وإظهار التأثيرات المتباينة لمختلف طرق التمثيل البياني.





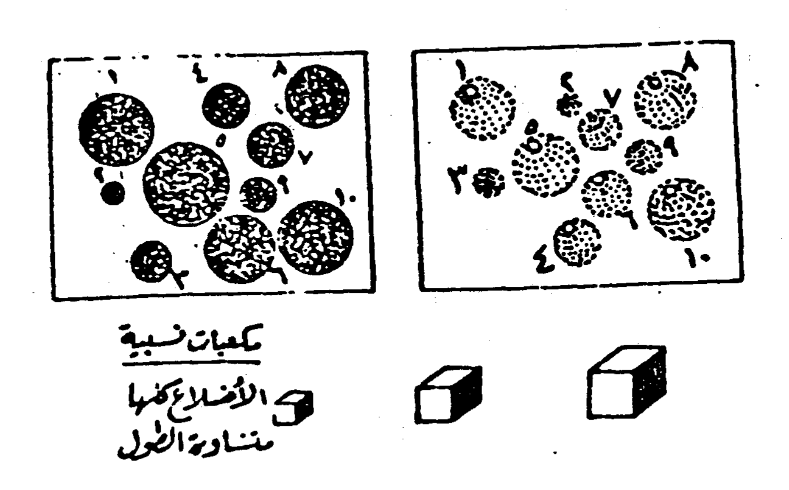
شكل رقم (٣ ـ ٩) الرسوم البيانية الدائرية

٤ ـ الرسوم البيانية الحجمية Three dimensional Graphs

إذا كانت البيانات المراد تمثيلها بيانياً ذات مدى عظيم جداً في القيم أو الكميات، فبدلاً من إدخال البعد الثاني (المساحية) للتغلب على مشكلة العظيمة التفاوت والاختلاف فإننا ندخل البعد الثالث الذي يترتب عليه استخدام رسوم بيانية حجمية تتناسب أحجامها مع مقدار الكميات التي تمثلها ومن أهم هذه الرسوم البيانية الكرات Spheres والمكعبات Cubes التي يقل استخدامها إلى حد كبير مثلها في ذلك مثل المربعات وعلى الرغم من مميزات هذا النوع من الرسوم البيانية فإن هناك بعض من المثالب التي يمكن إجمالها في: إن رسم الرسوم الحجميـة ليس أمراً سهلاً بل يتطل جهـداً وعملاً إضافياً حتى يبدو الشكـل الحجمي واضحاً، أو بمعنى آخر أن نعطي الكـرة أو الملعب الشكل الحجمي بإبعاده الثلاثة على سطح لوحة الرسم المستوى. وعلى الرغم من أن العلاقة بين أحجام الأشكال والكميات التي تمثلها صحيحة رياضياً إلا أنه ليس من السهل تقدير أحجام هذه الأشكال بمجـرد النظر إليها عكس الأعمدة البيانيـة. كذلك وعلى عكس الرسوم الدائرية التي يمكن تقسيمها لبيان تفصيلات الظاهرة، إلا أن الرسوم الحجمية لا يمكن تقسيمها لتوضيح أي بيانات تفصيلية وهـذا من أهم عيوب استخدام الأشكال الحجمية كالكرات والمكعبات.

وفي حالة استخدام الرسوم البيانية الحجمية الكرات والمكعبات لتمثيل كميات عظيمة التفاوت والاختلاف فإن حجم هذه الأشكال تتناسب مع مكعب نصف القطر (في حالة الكرات) أو مع مكعب طول الضلع (في حالة المكعبات) فالكرة الأكبر عشرة مرات من كرة أخرى سوف تمثل كمية أكبر ألف مرة $(١٠)^٣$ من الكمية التي تمثلها الكرة الأخرى. وكما هو متبع في طريقة رسم الدوائر البيانية، فإننا نستخرج أولاً الجذور التكعيبية للكميات، ونعتبر هذه الجذور التكعيبية أنصاف أقطار للدوائر التي سنعطيها شكل الكرات، أو نعتبرها طول ضلع المكعبات المراد رسمها. وفي حالة رسم الكرة نبدأ أولاً برسم دائرة عادية ثم

نعطيها الشكل الحجمي، أما أن نجعلها تمثل شكل «الكرة الأرضية» وذلك برسم شبكة رمزية من دوائر العرض وخطوط الطول فوق الدائرة المنرغة والتي ستبدد في النهاية على شكل كرة مجسمة، وأما أن نطمس كل مساحة الدائرة باللون الأسود مع ترك مساحة بيضاء في أعلى الكرة بحيث تبدو كالنور الساطع في أعلى الكرة شكل رقم (٣ ـ ١٠) أما المكعبات فهناك نوعان منها: نوع يبدو على شكل الدولاب وفيه يكون طول الجوانب مساوية لنصف طول الوجهة. والنوع الآخر يبدو متساوي الأضلاع والارتفاعات وتكون أطوال الجوانب والوجهة متساوية. ويعد أحسن شكل للمكعب هو الذي يكون فيه طول ضلع جوانبه ـــ طول ضلع واجهته، بحيث تميل هذه الجوانب من ٣٠ إلى ٥٠ من الخط الأفقي وتكون جوانب المكعب على يمين الناظر للرسم البياني.



(شكل رقم: ٣ ـ ١٠ أ) الرسوم البيانية الحجمية (الكرات والمكعبات)

وكمثال يمكن تمثيل عدد سكان كل من القاهرة والإسكندرية والجيزة (أكبر المدن المصرية) بالكرات والمكعبات كما هي الحال في الجدول التالي والشكل رقم (٣ ـ ١٠ ب).

عدد سكان أكبر المدن المصرية (١٩٦٦)

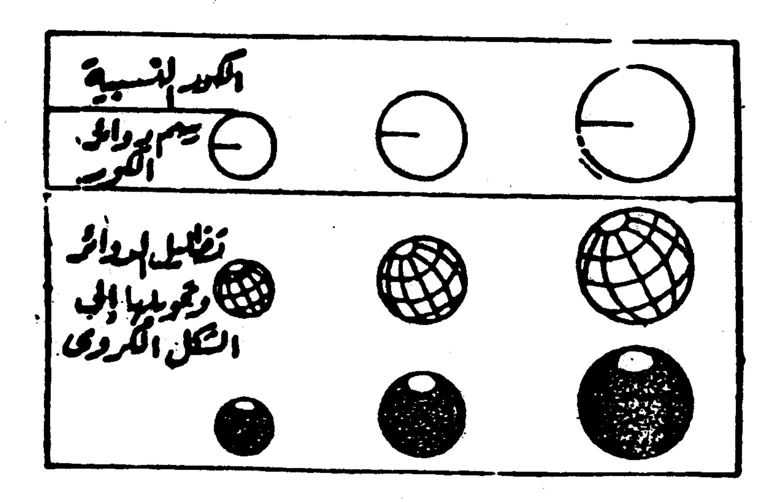
عدد السكان المدينة	عدد السكان (بالآلاف)	الجذر التكعيبي	طول قطر الكرة أو ضلع المكعب	الجذر التكعيبي / ٢٠٠
القاهرة	٤٫٢٢٠	١٦١٫٦	٠٫٨	سنتيمتر
الإسكندرية	١٫٨٠١	١٢١٫٦٤	٠٫٦	سنتيمتر
الجيزة	٥٧٠	٨٢٫٩٧	٠٫٤	سنتيمتر

٥ ـ الرسوم البيانية المثلثية

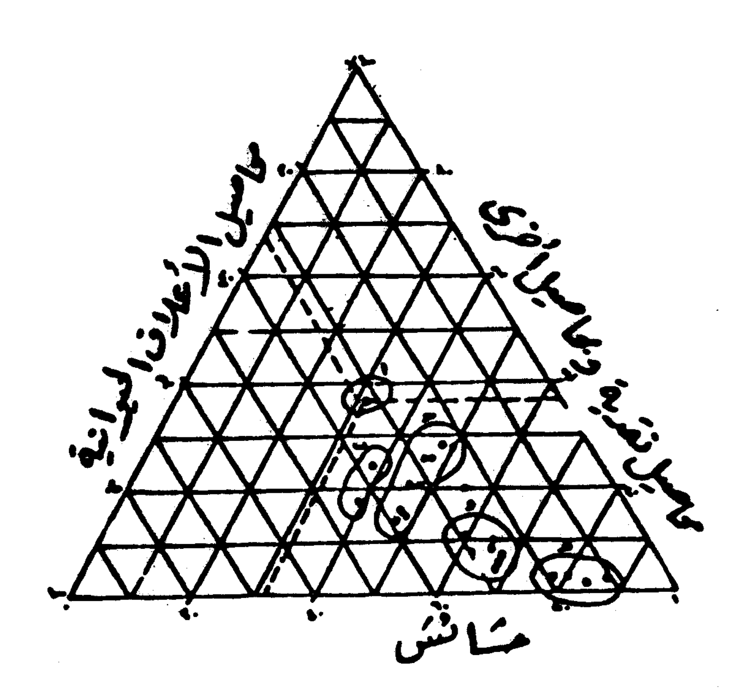
تستخدم الرسوم البيانية المثلثية في تمثيل البيانات النسبية الخاصة بثلاث ظاهرات مختلفة أو البيانات الأساسية الخاصة بثلاث ظاهرات مختلفة أو البيانات الأساسية الخاصة بثلاثة عناصر لظاهرة واحدة (مثل بيانات العمالة في المصانع، أنراع الحيوانات، أنواع المحاصيل، نباتات... تحليل التربة وبعض عيانتها) وذلك لمعرفة النسبة الغالبة بين الظاهرات أو الصفة السائدة بين عناصر الظاهرة بوجه عام.

وتقوم فكرة هذه الرسوم على أساس رسم مثلث متساوي الأضلاع يقسم كل ضلع منه إلى عشرة أقسام متساوية تستخدم كمقياس نسبي يبدأ من الصفر حتى

١٠٠٪ ويكون التقسيم في اتجاه عقرب الساعة أو بمعنى آخر أن يكون رقم ١٠٠ على أحد الأضلاع هو رقم الصفر للضلع المجاور والعكس مع رقم الصفر فيكون رقم ١٠٠ للضلع المجاور. وهكذا.. وبعد ذلك فصل بين كل رقم على أحد الأضلاع والرقم على الضلع المجاور ليكون مجموع هذين الرقمين ١٠٠ وبذلك نحصل على مجموعة من المثلثات الداخلية كل منها يشابه المقياس الكبير، ومنها نجد أن مجموع النسب لثلاثة عناصر إذ أضيفت لبعضها لنحصل على الرقم ١٠٠ يمكن تمثيله على الرسم البياني المثلثي بنقطة واحدة فقط. وفي عملية توقيع مكان هذه النقطة نبحث أولاً عن القيمة المراد تمثيلها على أحد الأضاع التي تمثل إحدى الظاهرات أو أحد العناصر ونمد منها خطاً يلتقي مع الخط الذي يعد من مكان القيمة الثابتة على أحد الضلعين الأخرين. ونقطة تلاقي الخطين هو موقع النقطة التي ستمثل الثلاث ظاهرات ويلتقي حتماً الخط الواصل من مكان القيمة الثالثة على الضلع الثالث إلى النقطة التي ستجمع الثلاث قيم في موضع واحد على الرسم (شكل رقم ٣ ـ ١١).



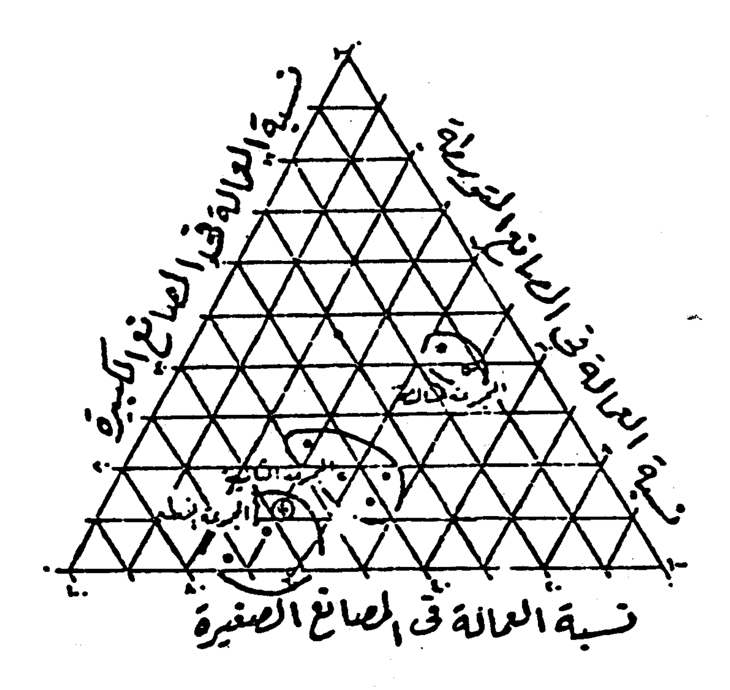
شكل رقم (٣ ـ ١٠ ب) الكرات النسبية



شكل رقم (٣ ـ ١١) رسم بياني مثلثي يبين الاستخدام
الغالب للأراضي الزراعية في السويد

ومن الأمثلة الشائعة لاستخدام هذه الرسوم ما يختص بتحليل التربة وعيناتها. فمثلاً إذا كانت لدينا مجموعة من عينات التربة وكان تحليلها على أساس النسب المئوية للعناصر الثلاثة الرئيسية التي تتألف منها وهي الرمل، الغرين، الصلصال وكان المطلوب تمثيلها بيانياً لمعرفة الصفة الغالبة للتربة بوجه عام كان من الممكن عندئذٍ استخدام هذا النوع من الرسوم البيانية وبالمثل يمكن استخدامه لبيان الحالة العامة لثلاث من أنواع الصناعات في مجموعة من المدن. ولهذه أهمية خاصة إذ أنه يمكن أن تتخذ كأساس لوضع تصنيف أو استخدام أنماط العديد من الظاهرات عن طريق تحديد بعض المساحات على الرسم والتي منها يمكن أن نعرف أي موقع للقيمة الثلاثية بالقرب من أحد أركان (نقط المثلث) يعني أن قيمة أحد العناصر (إحدى الظاهرات) لا بد أن تكون كبيرة جداً، بينما وقوع القيمة الثلاثية بالقرب من جوانب المثلث يشير إلى أن قيمة أحد العناصر (إحدى الظاهرات) لا بد أن تكون صغيرة جداً.

وكمثال تطبيقي يمكن سرده لبيان استخدام هذا النوع من الرسوم البيانية، استخدمت بيانات العمالة في ١١ مصنعاً رئيسياً بالسويد. وقسمت هذه المصانع حسب حجم العمالة المدربة بها (١٠٠ ـ ٥٠٠ عامل) إلى ثلاث فئات هي: مصانع صغيرة مصانع متوسطة الحجم، مصانع كبيرة، والشكل رقم (٣ ـ ١٢). يوضح التمثيل البيانات للفئات الثلاث من المصانع، ومنه يمكن أن نرى، أن الصناعات السويدية يمكن أن تصنف إلى ثلاث مجموعات (بدون الصناعات الهندسية والحديدية التي تقف كمجموعة بمفردها). المجموعة الأولى تشتمل على التحجير، الطباعة، الأعمال الخشبية، الصناعات الغذائية والمشروبات. وصناعات هذه المجموعة تقوم بالتدريب فيها المصانع الصغيرة، بينما لا توجد المصانع الكبيرة في مجال صناعات هذه المجموعة. والمجموعة الثانية تشمل الصناعات الجلدية، الغاز. المياه والكهرباء، الصناعات الكيميائية، والنسيج. وهي صناعات لا يتعادل التدريب في إنتاجها من الأنواع الثلاثة من المصانع حيث نجد أن هناك نسبة عمالة مدربة صغيرة للمصنع الكبيرة بينما تتعادل تقريباً نسبة العمال التي



شكل رقم (٣ ـ ١٢) رسم بياني مثلثي لتحديد فئات المصانع ونسبة العمال بها لعدد ١١ مصنعاً بالسويد

تدربها المصانع المتوسطة الحجم والمصانع الصغيرة التي تهتم بأنشطة المجموعة الصناعية الثانية. أما المجموعة الثالثة فهي الصناعات التعدينية، تصنيع الورق، وهذه تتساوى فيها نسبة العمالة المدربة تقريباً في المصانع الكبيرة والمتوسطة الحجم، بينما تقل نسبة العمالة المدربة كثيراً لإنتاج صناعات هذه المجموعة في المصانع الصغيرة الحجم.

٦ ـ أهرامات السكان Phpulation Pyramids

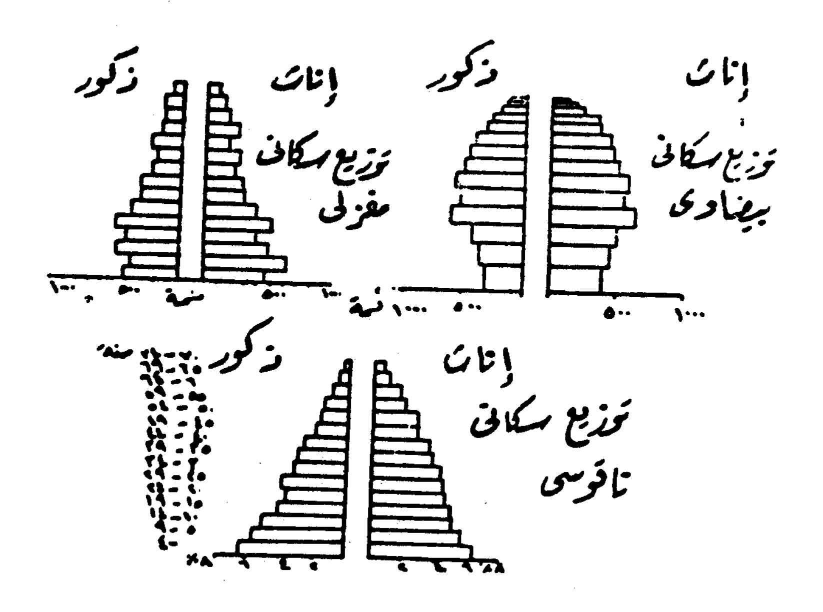
تستخدم الأهرامات السكانية كأحد طرق التمثيل البياني للبيانات الديموجرافية وبصفة خاصة لبيانات التركيب النوعي والعمري للسكان، حيث يجمع الهرم السكاني نسب كل من الذكور والإناث إلى العدد الكلي للسكان للفئات العمرية المختلفة. والهرم السكاني عبارة عن أعمدة بيانية أفقية ترسم على محورين أفقين أحدهما يمثل أعداد (أو نسب السكان الذكور) والآخر يمثل أعداد (أو نسب الإناث)، أما المحور الرأسي لها فهو يمثل فئات العمر كل منهما. ويجب أن تقسم المحاور الأفقية بنفس المقياس سواء للذكور أو الإناث.

ومن المفيد استخدام هذا الأسلوب من التمثيل البياني في معرفة الخصائص وتشخيص الاتجاهات للمجتمعات السكانية وكذلك عمل المقارنات عن حالة السكان لأكثر من أقليم أو دولة وإظهار الصفات العامة للسكان باستخدام التعدادات السكانية لأقليم أو دولة. وهناك عدة أنواع من أهرامات السكان نوجزها فيما يلي: ـ

أ ـ الهرم السكاني البسيط.

وتقوم فكرة على الأساس السابق شرحه لإنشاء ورسم الهرم السكاني ويستخدم هذا النوع لبيان الصفات العامة لسكان دولة أو أقليم معين. ومن المعروف لدى علماء الديموجرافيا أن لكل دولة هرم سكاني يميز تركيبها السكاني من حيث النوع والعمر لتعداد معين. وبناء على ذلك فإن أشكال الأهرامات السكانية ستختلف باختلاف التركيب النوعي والعمري للسكان بين البلاد المختلفة وهذا الاختلاف هو الذي يبرز المميزات ويؤكد الاتجاهات السكانية التي بالتالي تعطي صورة واضحة عن التركيب العمري والنوع للمجتمعات السكانية لهذه الدول (شكل رقم ٣ ـ ١٣)، فمثلاً إذا كان الهرم السكاني يأخذ الشكل المغزلي المقلوب فإن ذلك يدل على أن المجتمع الذي يمثله يتميز بتعادل معدلات المواليد والوفيات

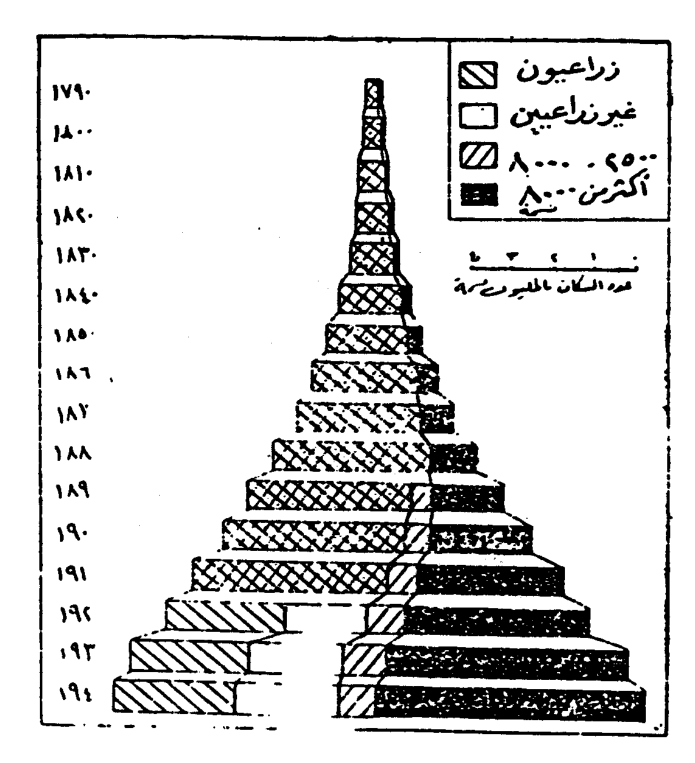
فيه   وإذا كان الهرم السكاني يأخذ الشكل البيضاوي من أعلى (أي في الفئات ذات الأعمار الكبيرة) وليس من عند القاعدة (أي في الفئات ذات الأعمار الصغيرة) فإنه يدل على أن المجتمع الذي يمثله مجتمعاً مسناً. ويستنتج ذلك من انخفاض نسبة الأطفال (ذكور وإناث) وزيادة نسبة المسنين (للذكور والإناث). أما إذا كان الهرم السكاني يتخذ شكلاً قريب من شكل الناقوس (الجرس) حيث تكون قاعدته عريضة ومحدب بلطف فإن ذلك يدل على ارتفاع معدلات الحضرية.



شكل رقم (٣ ـ ١٣) الأهرامات البيانية البسيطة

ب ـ الأهرامات السكانية المركبة Compound Pyramids

وتقوم فكرة هذا النوع من الأهرامات السكانية على أساس تمثيل التركيب النوعي أو العمري للسكان بأعمدة طول كل عمود يتناسب مع العدد الكلي للسكان لكل تعداد من التعدادات وبعد ذلك يقسم العمود (مثل طريقة الأعمدة المركبة) إلى أقسام الظاهرة الفرعية كان يقسم مثلاً إلى سكان الريف والحضر لكل تعداد. ففي الشكل رقم (٣ ـ ١٤) يمكن ملاحظة أنه خلال الفترة عن ١٧٩٠ إلى ١٨٨٠ كان هناك فئتين فقط بالإضافة إلى فئة من سكان يزيدون على ٨٠٠٠ نسمة والتي يمكن أن نعتبرها معبرة عن سكان الحضر. أما في الفترة من ١٨٩٠ إلى ١٩١٠ فإننا نلاحظ ثلاث تقسيمات عندما أدخلت فئة السكان أكثر من ٢٥٠٠ نسمة وأخيراً



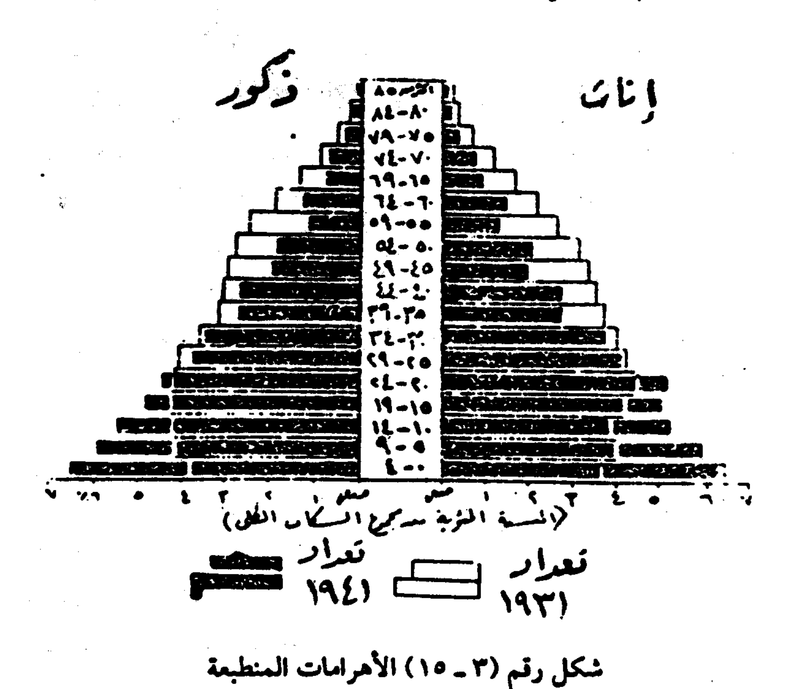
شكل رقم (٣ ـ ١٤): الهرم البياني المركب

فإنه من سنة ١٩٢٠ حتى ١٩٤٠ فإن أعمدة الهرم قد ارتفع عدد تقسيماتها الداخلية لتضم فئتين للمناطق الريفية الزراعية Rural Farm والمناطق الريفية غير الزراعية Rural Nonfarm.

جـ - الأهرامات السكانية المنطبعة Superimposed Pyramids

ويستخدم هذا النوع من الأهرامات لتمثيل بيانات التركيب النوعي والعمري في مكان ما لعدة تعدادات مختلفة وذلك بقصد المقارنة بين سكان كل تعداد وآخر، كما يمكن استخدام هذه الطريقة لمقارنة حالة السكان من حيث التركيب العمري والنوعي لسكانين للوقوف على مدى اختلاف توزيع السكان في أحدهما عن الآخرة.

وفي الحالتين يمكن رسم هرم سكاني بسيط بالطريقة السابقة ذكر وإعطائه لوناً أو ظلاً معيناً. ثم يرسم بعد ذلك هرماً بسيطاً أيضاً للسكان بنفس مقاييس الرسم المختلفة على الهرم السكاني الأول فيبدو وكأنه منطبعاً عليه (شكل رقم ٣ ـ ١٥).



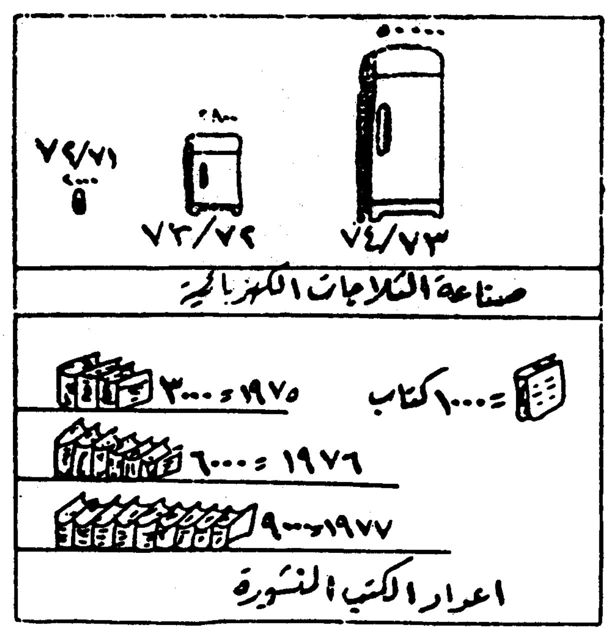
شكل رقم (٣ ـ ١٥) الأهرامات المنطبعة

ثانياً: طريقة التمثيل البياني بالرسوم التصويرية

تعتبر طريقة الرسوم البيانية التصويرية من أحسن وسائل الإيضاح وأكثرها جاذبية في التعبير عن تغير وتطور الظواهر. وتقوم فكرة هذه الطريقة على أساس إعطاء وحدات قياس الظاهرة أشكالاً تصويرية. فمثلاً عدد السكان يمكن أن يمثله برسم عدد من الأشخاص يمثل الشخص الواحد عدد مليون نسمة، أو عدد الثلاجات لأحد مصانع الثلاجات يمكن تمثيله برسم عدد من الثلاجات كل واحدة منها يمثل عشرة أو ما له ثلاجة، وكذلك عدد الكتب التي تصدرها إحدى دور النشر تصور بكتاب لكل عدد معين من هذه الكتب، كما أن عدد قراء إحدى الصحف اليومية يصور على أساس قارىء لكل عدد معين من الأعداد الصادرة (شكل رقم ٣ ـ ١٦). وفي كل الحالات السابقة يجب مراعاة الدقة في الرسم على أساس الحصول على عدد من الرسوم التوضيحية أو الرموز تشابه تماماً الظاهرة المراد تمثيلها. وعند تحديد عدد الوحدات من الظاهرة والتي تمثلها الشكل المختار فإنه يجب تحديد هذا العدد في ضوء أكبر وأصغر قيم في البيانات وكذلك في ضوء مساحة اللوحة المخصصة للرسم. ويمكن كذلك تكبير أو تصغير الوحدة التي تمثل الظاهرة ويراعى أن يكون هذا التكبير والتصغير على أساس مقياس الرسم لعدد المرات التي تساويها الوحدة الصغيرة. أما كسور القيم فيمكن تمثيلها برموز أو أشكال غير كاملة.

ويعاب على هذه الطريقة رغم أنها تشد أنظار الشخص العادي في التعرف على طبيعة الظاهرة من حيث تطورها وتغيرها إلا أنها لا تعطي فكرة دقيقة عن قيم الظاهرة حيث أنه نضطر إلى تقريب القيم إلى أقرب عشرة أو مائة أو ألف حتى يمكن التخلص من الكسور الصغيرة والتي لا يمكن أن نعطي لها شكلاً أو رمزاً. فمثلاً القيمة ٦٠٠ كتاب من الصعب تمثيلها تصويرياً برمز أو شكل متكامل إذا كان الأخير يمثل ١٠٠٠ كتاب مثلاً. كذلك لا تلائم هذه الطريقة تمثيل البيانات التي تتميز بتباين مفردات قيمها وأيضاً لا تعطي هذه الطريقة فكرة حقيقية واضحة عن حقيقة المقارنة بين هذه الوحدات.





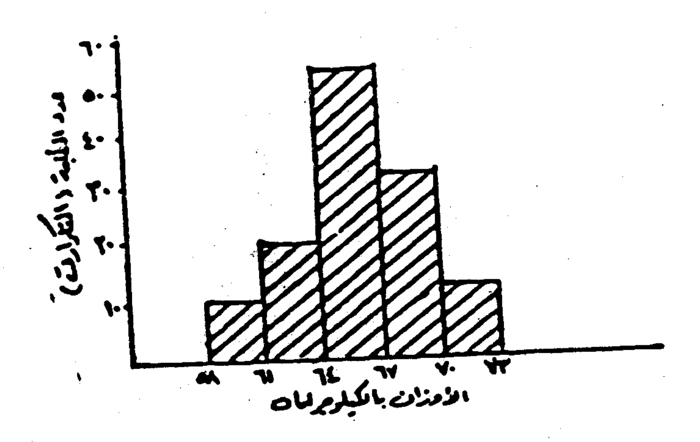
شكل رقم (٣ ـ ١٦) الرسوم البيانية التصويرية

طرق العرض البياني للبيانات التكرارية (المبوبة)

تمثل بيانات التوزيعات المبوبة (التكرارية) بعدة طرق بيانية مختلفة أهمها: المدرج التكراري، الضلع التكراري والمنحنى التكراري:

١ ـ المدرج التكراري Histogram

يتكون المدرج التكراري من مجموعة من المستطيلات المتلاصقة التي تكون قاعدتها على المحور الأفقي (محور السنات) وطول هذه القاعدة يساوي طول الفئة، كما تتناسب مساحة كل مستطيع مع التكرار المناظر لكل فئة على المحور الرأسي (شكل رقم ٣ ـ ١٧).



شكل رقم (٣ ـ ١٧) هيستوجرام يبين توزيع أوزان الطلبة

وإذا كانت الفئات كلها لها نفس الطول، فإنه من المعتاد أن تأخذ الارتفاعات مساوية لتكرارات الفئات. أما إذا كانت الفئات غير متساوية الطول فإن هذه الأطوال يجب أن تعدل قبل رسم المدرج التكراري. وذلك بقسمة تكرار كل فئة على طول هذه الفئة وبذلك نحصل على التكرار المعدل ثم ننشىء المدرج

التكراري برسم عدد من المستطيلات قاعدتها تمثل أطوال الفئات (غير متساوية) على المحور الأفقي وارتفاعاتها هي التكرارات المعدلة. وفي هذه الحالة تتناسب مساحة كل مستطيل مع التكرارات المعتدلة المناظرة لكل فئة.

ومن الواضح أن المدرج التكراري يصلح لتمثيل المتغيرات المتصلة ولا يصلح لتمثيل المتغيرات غير المتصلة (الوثابة).

المضلع التكراري Frequency Polygon

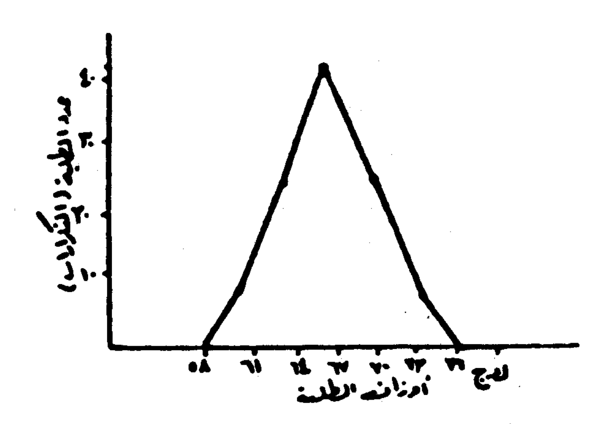
يمكن اعتبار إنشاء المدرج التكراري بالطريقة التي ذكرناها سابقاً خطوة من خطوات رسم المضلع التكراري للتوزيع التكراري. وتعتمد فكرة إنشاء المضلع التكراري على فكرة التمثيل البياني من خلال الخط البياني حيث تبين فقط التمثيل تكرار الفئة المقابلة لمركز الفئة. ويرسم المضلع التكراري بإيصال نقط تنصيف رؤوس المستطيلات المكونة للمدرج التكراري بمجموعة من الأضلاع كل ضلع بين مركز فئة ما ومركز الفئة التالية مباشرة، وحتى يمكن قفل شكل المضلع التكراري وتحديد مساحته فإننا نفترض وجود فئتين قبل الفئة الأولى وبعد الفئة الأخيرة والتكرار المناظر لكل منها يساوي صفر، ثم يتم توصيلهما بأضلاع مع مركز الفئة الأولى والأخيرة فنحصل بذلك على شكل كامل للمضلع التكراري.

وفكرة استخدام مركز الفئة أو منتصفها في رسم المضلع التكراري يعتمد على افتراض تركز القيم عند متوسطها الحسابي حيث أن مركز الفئة يساوي:

$$\frac{\text{الحد الأدنى للفئة + الحد الأعلى للفئة}}{٢}$$

ويتميز المضلع التكراري بأن المساحة المحصورة تحت المضلع هي نفسها المساحة المحصورة تحت المدرج التكراري، ولكنه يكون أكثر دقة من المدرج التكراري من حيث إعطائه صورة أكثر واقعية لاتجاهات وخصائص التوزيع. ويعتبر المضلع التكراري من أنسب الطرق البيانية لتمثيل أكثر من توزيع واحد من

التوزيعات التكرارية مثلاً. والشكل التالي يوضح المضلع التكراري لأوزان الطلبة (جدول رقم: ٣ ـ ٤).



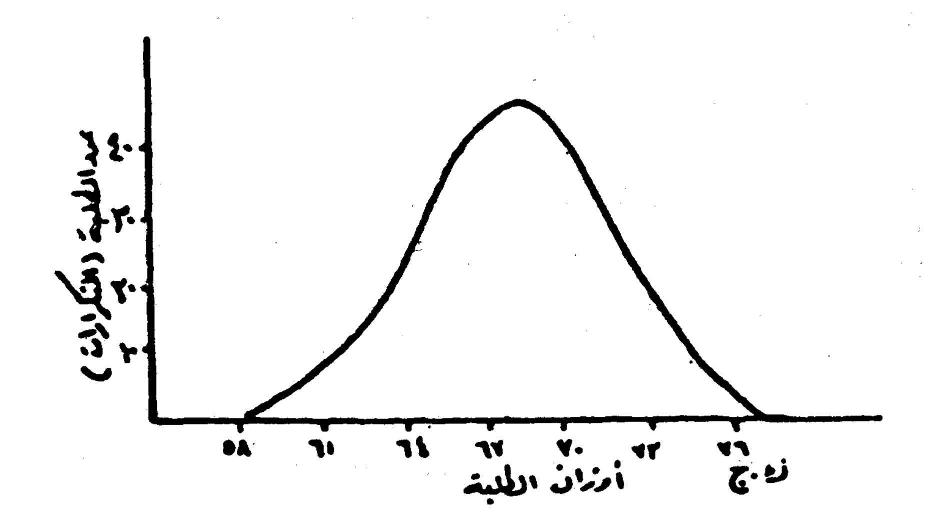
شكل رقم (٣ ـ ١٨) المضلع التكراري لأوزان الطلاب

المنحنى التكراري Frequency Curve

يمكن تمثيل التوزيعات التكرارية بطريقة أخرى تظهر في شكل هندسي واضح وذلك برسم المنحنى التكراري للتوزيع والذي نحصل عليه من خلال رسم المضلع التكراري أولاً وتمهيد خطوط المضلع المنكسرة.

ولرسم المنحنى التكراري نعين مراكز الفئات على حب التكرارات المناظرة ونوقع نقط المضلع التكراري ونمهد الخطوط المنكسرة بين هذه النقط باليد (توجد طرقاً رياضية لتكوين المنحنى التكراري تجعل المساحة تحت المنحنى مساوية

للمساحة تحت المضلع التكراري). بحيث يمر المنحنى بأغلبية رؤوس المضلع التكراري.



شكل رقم (٣ ـ ١٩) المنحنى التكراري لأوزان الطلاب

وفي الواقع كلما كان حجم التوزيع التكراري كبيراً وأطوال الفئات المستخدمة في التوزيع قصيرة كلما أدى ذلك إلى التوصل إلى منحنى تكراري أكثر دقة في تحديد خصائص واتجاهات التوزيع مما لو كان التوزيع صغير الحجم وطول الفئة كبيراً.

ونظراً لأن المنحنى التكراري يمكن رسمه من نقط المضلع التكراري، لذا فإن شكل المنحنى يتوقف على توزيع البيانات، ومن المنطقي أن نتوقع وجود عدد كبير من الأشكال المختلفة للمنحنيات التكرارية إلا أنه يمكن حصر أشكال المنحنيات التي تقابلنا عادة في التحليل الإحصائي للبيانات فيما يلي:-

١ ـ المنحنى التكراري المتماثل ذو الشكل الناقوسي الذي يتميز بأن المشاهدات المتساوية البعد عن مركز النهاية العظمى لها نفس التكرارات، ومن الأمثلة الهامة له المنحنى المعتدل.

٢ ـ المنحنيات التكرارية الملتوية والتي تتميز بأن أحد طرفيها يمتد أكثر من الآخر على جانبي مركز النهاية العظمى إذا كان الطرف الأيمن أطول فيكون المنحنى في هذه الحالة ملتوياً إلى اليمين أو ملتوياً التواء موجباً، بينما لو كان العكس صحيحاً بأن المنحنى يكون ملتوياً إلى اليسار أو ملتوياً التواء سالباً.

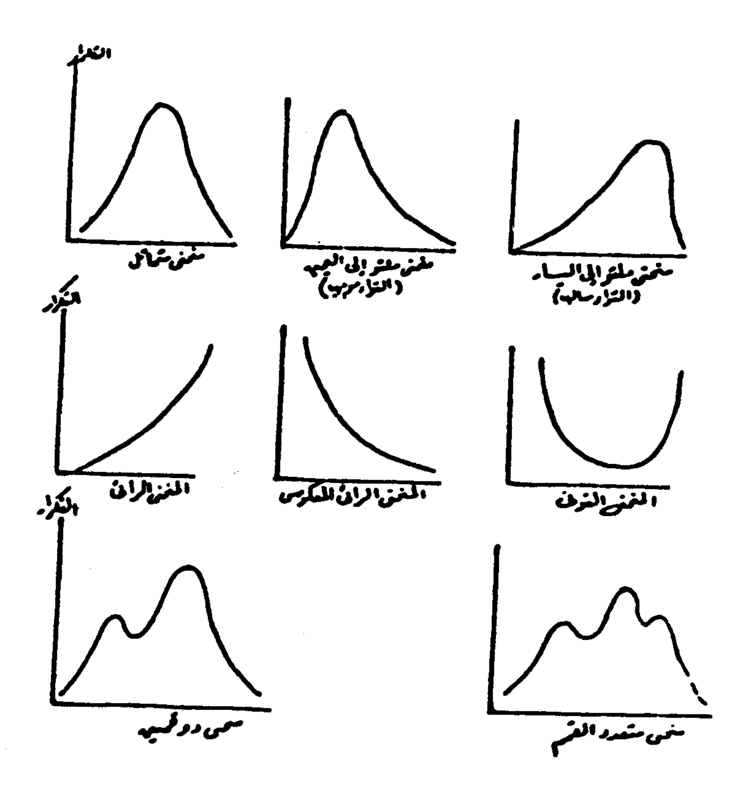
٣ ـ المنحنيات ذات الشكل الرائي أو الشكل الرائي المعكوس وفيها تقع نقطة النهاية العظمى للمنحنى عند أحد طرفي المنحنى.

٤ ـ المنحنى النوني الذي يتميز بأن له نهاية عظمى عند كل من طرفيه.

٥ ـ المنحنى ذو القمتين الذي يتميز بأن له نهايتان عظميتان.

٦ ـ المنحنى متعدد القمم والذي له أكثر من نهايتين عظميتين.

والشكل التالي يوضح أنواع المنحنيات السابقة:



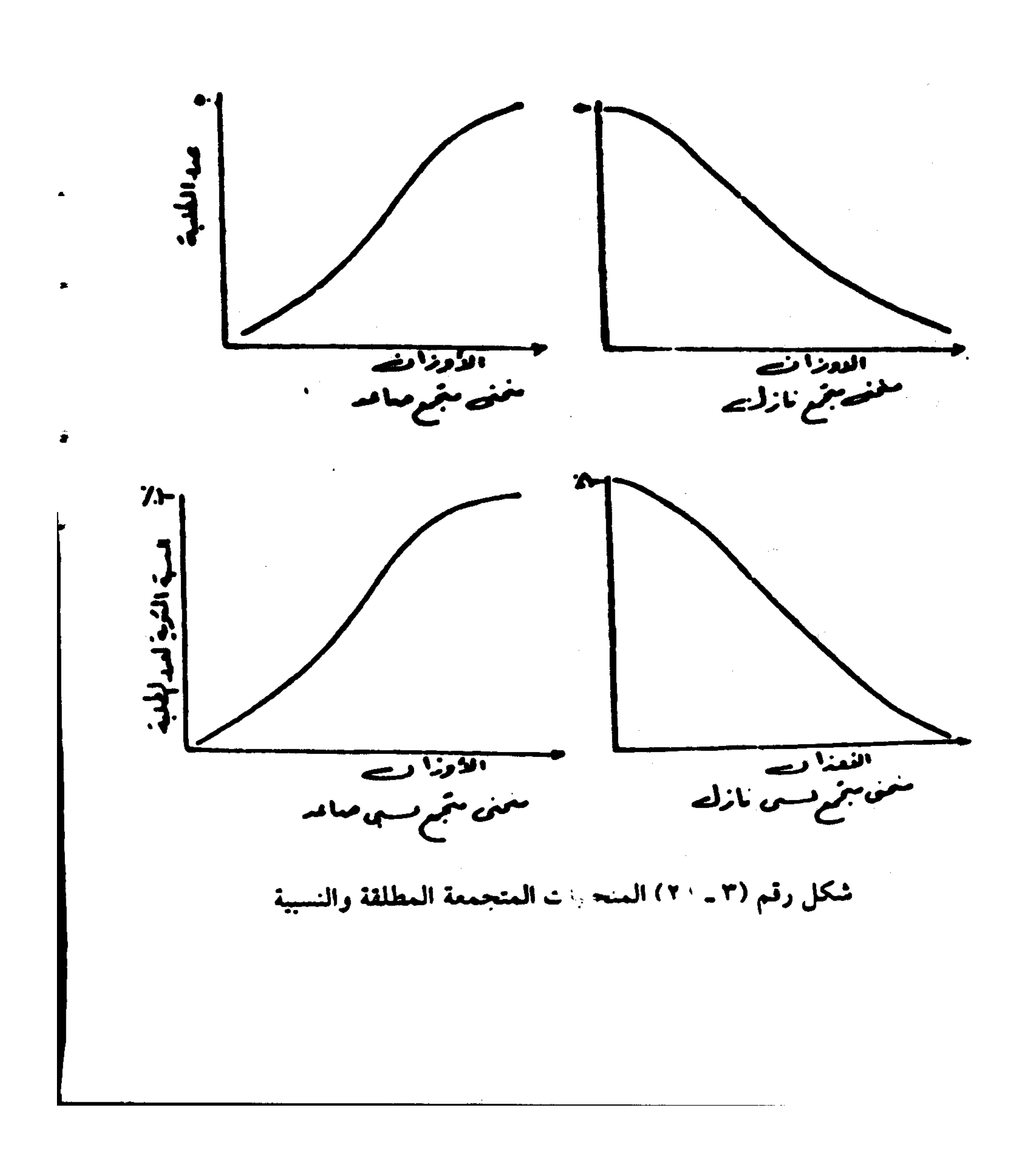
شكل رقم (٣ ـ ٢٠) أشكال المنحنيات البيانية

وإلى جانب هذه الأنواع مس المنحنيات التكرارية هناك أيضاً منحنيات بيانية تمثل جداول التوزيعات التكرارية المتجمعة

وقد سبق أن أوضحنا كيفية عمل الجداول التكرارات المتجمعة، الصاعدة أو النازلة، ولتمثيل هذه الجداول التكرارية بمنحنى يمكن رسم منحنى متجمع صاعد وكذلك منحنى متجمع نارل ڡفي الحاله الأولى نؤحد الحدود العليا للفئات على

المحور الأفقي والتكرارات المتجمعة على المحور الرأسي ونوقع النقط حسب إحداثيها الأفقي والرأسي ونصل بينها بمنحنى ممهد فنحصل على المنحنى المتجمع الصاعد، لأن التكرارات المتراكمة تكون في تزايد، أما في الحالة الثانية تؤخذ الحدود الدنيا للفئات الأولية على المحور الأفقي والتكرارات المتجمعة النازلة على المحور الرأسي ثم نصل بين النقط بمنحنى ممهد فنحصل على المنحنى المتجمع النازل.

والأشكال الآتية تبين المنحنيات المتجمعة والنسبية.



شكل رقم (٣ ـ ٢١) المنحنيات المتجمعة المطلقة والنسبية

الباب الثالث

المقاييس الإحصائية للبيانات

المقاييس الاحصائية للبيانات

بعد أن حصلنا على القيم المختلفة لجميع المفـردات محـل الدراسـة للظاهرة محل البحث وسواء كنا ندرس مجتمع الظاهرة بالكامل أو كان ذلـك يتم من خلال عينة من المجتمع، وسواء احتفظنا بتلك القيم كمـا هـى ودون إجراء أى عمليات تبويب لها أو سواء تم تنظيم وتبويب تلك البيانات فى توزيع تكرارى فإنه يصبح لزاما علينـا معرفـة مختلـف المؤشـرات والمقـاييس الإحصائية الأساسية التى يمكن استخراجها من البيانات لتصف الظاهرة محـل الدراسة بحيث تمثل تلك المقاييس الاحصائيـة القـاعدة الأساسـية للتحليـل الأحصائى للبيانات.

ويمكن القول بأن السمات الأساسية لمعظم الظواهر الكمية هى:

١- وجود قيمة متوسطة تتجمع وتتركز حولها معظم مفردات الظاهرة بحيث يمكن استخدام تلك القيمة المتوسطة ملخصا ممثلا معـبرا عـن جميـع قياسات ومفردات الظاهرة وهو مانطلق عليه النزعة المركزية للبيانـات أو المتوسط.

٢- تشتت وانتشار مختلف مفردات الظاهرة حول المتوسط بنمط واضح يعبر عن اختلافات المفردات فيما بينها وبين المتوسط.

٣- الشكل العام للمنحنى التكرارى المناظر للتوزيـع التكـرارى لمفـردات الظاهرة ويدخل فى ذلك

ا) نوعية اتجاه الالتواء إن وجد فى المنحنى التكرارى للقيم

ب) عدد قمم المنحنى التكرارى ودرجةالتفرطح أو التدبب للمنحنى وحيدالقمة

وينمثل نحليلنا لمعرفات الظاهرة فى إيجاد المقاييس الكمية المناسبة للسمات السابقة مما يمكن معه إيجاد وصف تقريبى للظاهرة من خلال عدد من المقاييس الوصفية التى يمكن تفسير معناها. ومن المؤكد أن هذه المقاييس الوصفية تمثل المعلومات الأساسية التى يجب معرفتها عن الخصائص الرئيسية للظاهرة وتمثل المعلومة الاحصائية الأساسية فى التحليل الاحصائى والذى يمكن إستخدامه فى اتخاذ القرارات.

وبشىء من التفصيل نبدأ فى مناقشة مختلف المؤشرات والمقاييس الإحصائية التى تمثل الخطوة الأساسية الأولى فى تحليل ومعرفة خصائص المجتمعات الاحصائية والبيانات.

أولا : المتوسط (النزعة المركزية):

أ – الوسط الحسابى:

يعتبر الوسط الحسابى أكثر المتوسطات شهرة واستخداما للتعبير عن الظواهر الكمية وإن كان فى بعض الأحيان يكون من الأفضل عدم استخدامه كما أن هناك نوعا معينا من البيانات يجب فيها استخدام متوسطا آخر محددا ليكون هذا المتوسط معبرا بصورة دقيقة عن تلك البيانات ويمكن تعريف الوسط الحسابى جبريا بأنه مجموع المشاهدات مقسوما على عددها. فإذا كانت لدينا ن من المشاهدات هى:

$$س_١ ، س_٢ ، س_٣ ، \dots ، س_ن$$

فإن الوسط الحسابى وسنرمز له بالرمز $\bar{س}$ يكون:

$$\bar{س} = \frac{مجـ س}{ن} \qquad (١)$$

ونلاحظ أن هذه هى الصيغة الرياضية للوسط الحسابى للبيانات غير المبوبة ومن البديهى أن هناك صيغة أخرى للبيانات المبوبة سنناقشها بعد قليل:

مثال (١)

توضح القيم التالية الدخل الشهرى لعينة عشوائية من عشرة من العـــاملين فى جامعة عين شمس:

٢٠٠ ٦٠٠ ٣٢٥ ١٨٠ ٢٤٦

٥٣٠ ١٦٤ ٣٨٠ ٤٥٠ ٣٧٥

والمطلوب إيجاد الوسط الحسابى باعتباره المتوسط الذى يعبر عن المتوســـط الشهرى لدخل الفرد الواحد فى تلك العينة.

الحــــل

الوسط الحسابى $= \bar{س} = \frac{مجـ س}{ن}$

$$= \frac{٢٠٠ + ٦٠٠ + ٣٢٥ + ١٨٠ + ٢٤٦ + ٥٣٠ + ١٦٤ + ٣٨٠ + ٤٥٠ + ٣٧٥}{١٠}$$

$$= \frac{٣٤٥٠}{١٠} = ٣٤٥$$

ونلاحظ مايلى:

١- يتسم الوسط الحسابى بالبساطة والسهولة عند إيجاد قيمته حسابيا.

٢- تم استخدام جميع مفردات العينة (عشرة مفردات فى هذا المثـــال) عنـــد إيجاد قيمة الوسط الحسابى ولاشك أن استخدام جميع المفــردات والقيـــم محل الاعتبار فى إيجاد قيمة الوسط الحسابى يؤدى الى إرتفـــاع كفـــاءة الوسط الحسابى بشكل عام كمتوسط للقيم.

٣- تم استخدام جميع القيم والمفردات مباشرة ودون إجراء أى عملية ترتيب تصاعدى أو تنازلى للقيم مما يمكن معه القول بأن الوسط الحسابى يعتبر مقياسا للقيم ذاتها وليس مقياسا لمنازل ورتب تلك القيم.

٤- نلاحظ أن قيمة الوسط الحسابى الناتجة فى المثال السابق ليست إحدى القيم المستخدمة فى العملية الحسابية ولكنها قيمة تتركز حولها وتتجمع مختلف قيم ومفردات العينة مما يعنى أن قيمة الوسط الحسابى قد تكون أو لاتكون إحدى قيم المشاهدات الداخلة فى العملية الحسابية ولكن قيمة الوسط الحسابى تمثل المتوسط الذى يعبر عن المفردات والمشاهدات ككل.

٥- يمكن القول أن مجموع انحرافات القيم عن وسطها الحسابى يساوى الصفر وهذه خاصية رياضية من خصائص الوسط الحسابى، أى أن:

مجـ (س − $\overline{س}$) = صفر

وبتطبيق ذلك على قيم المثال السابق نجد أن:

مجـ (س − $\overline{س}$) = (٢٠٠ − ٣٤٥) + (٦٠٠ − ٣٤٥) + (٣٢٥ - ٣٤٥)

+ (١٨٠ − ٣٤٥) + (٢٤٦ − ٣٤٥) + (٥٣٠ − ٣٤٥) + (١٦٤ − ٣٤٥)

+ (١٦٤ − ٣٤٥) + (٣٨٠ − ٣٤٥) + (٤٥٠ − ٣٤٥) + (٣٧٥ − ٣٤٥)

= − ١٤٥ + ٢٥٥ − ٢٠ − ١٦٥ − ٩٩ + ١٨٥ − ١٨١ + ٣٥ + ١٠٥

+ ٣٠ = − ٦١٠ + ٦١٠ = صفر

ومن هذه الخاصية للوسط الحسابى يمكن أن نلاحظ أن:

مجـ س = ن $\overline{س}$

أى أن مجموع القيم = عدد القيم × الوسط الحسابى للقيم.

نلاحظ أيضا أن الوسط الحسابى يتأثر بشدة بالقيم المنضرفة، ففى المثال السابق بفرض أن هناك عاملا إضافيا تم اختياره عشوائيا ليصبح عدد مفردات العينة = ١١ وقد كان الدخل الشهرى لهذا العامل الاضافى ١٤٤٥ نلاحظ أن الوسط الحسابى الجديد يصبح.

$$\bar{س} = \frac{مجـ س}{ن}$$

$$= \frac{٤٨٩٥}{١١} = ٤٤٥ \text{ جنيها}$$

مما يعنى أن وجود وإضافة مفردة متطرفة واحدة تقفز بالمتوسط الشهرى للدخل من ٣٤٥ جنيه الى ٤٤٥ جنيه ويصبح الوسط الحسابى فى حالة وجود مثل هذه القيمة المتطرفة غير معبرا بصورة جيدة عن متوسط قيم الظاهرة.

الوسط الحسابى المرجح:

يحدث فى كثير من الأحيان أن تتكرر بعض قيم مفردات الظاهرة وبالتالى يتوفر لدينا بيانا بقيم الظاهرة مع وجود عدد مرات التكرار لكل قيمة ولهذا يجب مراعاة ذلك عند إيجاد قيمة الوسط الحسابى لهذه القيم المرجحة

مثال (٢)

يوضح البيان التالى الدخل الشهرى لعينة عشوائية من ١٥ موظفا بجامعة عين شمس

الدخل الشهرى للموظف	عدد الموظفين
٢٠٠	٣
٢٨٠	٦
٣٢٠	٤
٣٦٥	٢
المجموع	١٥

والمطلوب إيجاد الوسط الحسابى للدخل الشهرى للموظف فى هذه العينة:

الحــــل

نلاحظ أن البيان السابق يعنى أن قيم مفردات العينة هى:

(٢٠٠، ٢٠٠، ٢٠٠)، (٢٨٠، ٢٨٠، ٢٨٠، ٢٨٠، ٢٨٠، ٢٨٠)، (٣٢٠، ٣٢٠، ٣٢٠، ٣٢٠)، (٣٦٥، ٣٦٥) ويكون حينئذ الوسط الحسابى:

$$\bar{س} = \frac{مجـ س}{ن}$$

$$= \frac{(٢٠٠ + ٢٠٠ + ...+ ٣٦٥ + ٣٦٥)}{١٥}$$

$$= \frac{٤٢٩٠}{١٥} = ٢٨٦ \text{ جنيها}$$

ويمكن أن نلاحظ أنه يمكن ضرب كل قيمة فى عدد مرات تكرارهـــا لإيجاد النتيجة السابقة، أى أن :

$$\bar{س} = \frac{(٢٠٠\times٣ + ٢٨٠\times٦ + ٣٢٠\times٤ + ٣٦٥\times٢)}{١٥} = ٢٨٦ \text{ جنيها}$$

وبصفة عامة اذا اعتبرنا أن كل قيمة من قيم الظاهرة س قد تكررت ك مـــن المرات أى أن لدينا

قيم الظاهرة : $س_١$ $س_٢$ .... $س_و$

عدد مرات تكرارها : $ك_١$ $ك_٢$ .... $ك_و$

فإن: الوسط الحسابى المرجح =

$$\bar{س} = \frac{(س_١ ك_١ + س_٢ ك_٢ + ... س_و ك_و)}{(ك_١ + ك_٢ + ... ك_و)}$$

$$\bar{س} = \frac{مجـ س ك}{مجـ ك} \quad ....(٢)$$

مثال (٣)

يوضح البيان التالى عدد أفراد الأسرة لكل طالب فى عينة عشوائية مـــن ٥٠ طالب من كلية التجارة.

عدد أفراد الأسرة للطالب	عدد الطلبة
٣	٨
٤	١٥
٥	١٠
٦	٥
٧	١٠
٨	٢
المجموع	٥٠

أوجد الوسط الحسابى لعدد أفراد الأسرة لكل طالب فى العينة

الحـــ ـــس

يمكن تنظيم العمل الحسابى مباشرة لإيجاد الوسط الحسابى كما يلى:

عدد أفراد الأسرة للطالب (س)	عدد الطلبة (ك)	س ك
٣	٨	٢٤
٤	١٥	٦٠
٥	١٠	٥٠
٦	٥	٣٠
٧	١٠	٧٠
٨	٢	١٦
المجموع	٥٠	٢٥٠

$$\text{الوسط الحسابى} = \bar{\text{س}} = \frac{\text{مجـ س ك}}{\text{مجـ ك}} = \frac{٢٥٠}{٥٠} = ٥ \text{ أفراد}$$

أى أنه فى المتوسط تتكون أسرة الطالب فى العينة العشوائية التى تم اختيارهـا من خمسة أفراد.

الوسط الحسابى من جدول تكرارى:

يمكن إيجاد الوسط الحسابى من الجدول التكرارى بنفس الكيفية السابقة فى الوسط الحسابى المرجح وذلك باعتبار أن التكرار الموجود فى كل فئة إنما هو الثقل الذى يعبر عن عدد مرات تكرار قيمة الفئة ومن ثم يعكس الترجيح للفئة كما أنه يتم استخدام مركز كل فئة ليكون هو القيمة المتوسطة للفئة والذى يعبر عن قيمة الفئة ولاشك أن هذا هو التصور الطبيعى والمنطقى حيث أن كل فئة مدى متسع من القيم يبدأ من الحد الأدنى للفئة وينتهى بنهاية الفئة ، لهذا فانه من المنطقى اختيار مركز الفئة ليكون القيمة المتوسطة التى تعبر

عن الفئة وجدير بالذكر أنه من الأفضل كقاعدة عامة استخدام القيم الأصلية للظاهرة قبل تبويبها لأيجاد الوسط الحسابى للظاهرة بصورة دقيقة، كما أن الوسط الحسابى للظاهرة المحسوب من جدول تكرارى يكون تقريبا للوسط الحسابى المحسوب من قيم الظاهرة مباشرة، ويكون هذا التقريب جيدا كلما كانت فئات التوزيع التكرارى أكثر تفصيلا وكلما كانت أطوال الفئات صغيرة

أيضا من البديهى أن يكون الجدول التكرارى مغلق من البداية ومن النهاية وذلك حتى يمكن إيجاد مراكز الفئات لجميع فئات الجدول.

وبناء على ماسبق فان خطوات إيجاد الوسط الحسابى من جدول تكرارى مغلق هى كمايلى:-

١- إيجاد مركز كل فئة فى الجدول (س)

٢- ترجيح قيمة كل فئة وذلك بضرب س × ك

يكون الوسط الحسابى $\overline{س}$ هو

$$\overline{س} = \frac{مجـ\ س\ ك}{مجـ\ ك}$$

مثال (٤)

يوضح الجدول التالى التوزيع التكرارى لعينة عشوائية من ٨٠ طالبا حسب فئات الوزن المختلفة:

فئات الوزن	٥٠-	٦٠-	٧٠-	٨٠-	٩٠-١٠٠	المجموع
عدد الطلبة	٨	١٢	٤٠	١٥	٥	٨٠

والمطلوب ايجاد الوسط الحسابى

الحـــــل

بإيجاد مركز الفئة لكل فئة من فئات الجدول (س) وبترجيح مركز كل فئـة بضربه فى تكرار الفئة فإن:

فئات الوزن	عدد الطلبة (ك)	مركز الفئة (س)	س ك
٥٠-	٨	٥٥	٤٤٠
٦٠-	١٢	٦٥	٧٨٠
٧٠-	٤٠	٧٥	٣٠٠٠
٨٠-	١٥	٨٥	١٢٧٥
٩٠-١٠٠	٥	٩٥	٤٧٥
المجموع	٨٠		٥٩٧٠

ويكون الوسط الحسابى س:

$$\bar{س} = \frac{مجـ س ك}{مجـ ك}$$

$$\text{الوسط الحسابى للوزن} = \bar{س} = \frac{٥٩٧٠}{٨٠} = ٧٤,٦ \text{ كيلو جرام}$$

ومن البديهى أنه يمكن ايجاد الوسط الحسابى من الجدول التكرارى غير متسـاوى الفئات وبنفس الخطوات والكيفية السابقة ويتضح ذلك من المثال التالى:

مثال (٥)

يوضح الجدول التالى التوزيع التكرارى للدخل الشهرى لعينـــة عشـــوائية من ٨٠ موظف بجامعة عين شمس:

الدخل الشهرى	١٠٠-	١٥٠-	٢٠٠-	٣٠٠-	٤٠٠-٦٠٠	المجموع
عدد الموظفين	٥	٢٥	٣٠	١٥	٥	٨

والمطلوب ايجاد الوسط الحسابى للأجر من هذا التوزيع التكرارى.

الحـــــل

نلاحظ أن التوزيع التكرارى السابق توزيع غير متساوى الفئات ولكن كما أشرنا فإن هذا لايغير من فلسفة وكيفية إيجاد الوسط الحسابى ولهذا فإنه يتم إيجاد مراكز الفئات للدخل ثم تقوم بضرب مركز كل فئة فى التكرار المقابل لها ثم يتم تطبيق قانون الوسط الحسابى من المعادلة (٢).

فئات الوزن	عدد الطلبة (ك)	مركز الفئة (س)	س ك
١٠٠-	٥	١٢٥	٦٢٥
١٥٠-	٢٥	١٧٥	٤٣٧٥
٢٠٠-	٣٠	٢٥٠	٧٥٠٠
٣٠٠-	١٥	٣٥٠	٥٢٥٠
٤٠٠-٦٠٠	٥	٥٠٠	٢٥٠٠
المجموع	٨٠		٢٠٢٥٠

$$\bar{س} = \frac{مجـ\ س\ ك}{مجـ\ ك} = \frac{٢٠٢٥٠}{٨٠} = ٢٥٣,١٢٥\ \text{جنيه}$$

الوسط الحسابى للدخل الشهرى:

ولكن فى بعض الأحيان قد تكون مراكز الفئات قيما كبيرة أو بها كسور فيكون من المجهد حسابيا الاستخدام المباشر لمراكز الفئات ويمكن حينئذ تبسيط العمل الحسابى بطرح وسط فرضى من جميع مراكز الفئات وبذلك نحصل على انحرافات مراكز الفئات عن الوسط الحسابى وسنرمز لها بالرمز (ح) ثم يتم ترجيح هذه الانحرافات بتكرارات الفئات ويصبح الوسط الحسابى فى هذه الحالة:

$$\text{الوسط الحسابى} = \text{الوسط الفرضى} + \frac{\text{مجـ ح ك}}{\text{مجـ ك}} \quad \ldots\ldots\ldots\ldots\ldots(٣)$$

مثال (٦)

يمثل الجدول التالى التوزيع التكرارى لعينة عشـــوائية مـــن ٨٠ موظفـــا بجامعة عين شمس حسب فئات العمر المختلفة.

العمر	٣٠-	٣٥-	٤٠-	٤٥-	٥٠-٥٥	المجموع
عدد الموظفين	٨	٢٢	٣٠	٢٤	٦	٨٠

والمطلوب إيجاد الوسط الحسابى للعمر بإستخدام طريقة الانحراف عن وســـط فرضى لمراكز الفئات.

الحـــــل

بإيجاد مراكز الفئات واختيار ٤٢,٥ كوسط فرضى يتم طرحه من جميــــع مراكز الفئات بهدف تبسيطها وبذلك نحصل على انحرافات مراكز الفئات عـن الوسط الفرضى (ح) وبترجيح الانحراف لكل فئة بتكرار تلك الفئة يمكـــن ان ننظم ذلك فى الجدول التالى:

العمر	ك	مراكز الفئات س	الانحرافات ح	ح ك
٣٠-	٨	٣٢.٥	-١٠	-٨٠
٣٥-	٢٢	٣٧,٥	-٥	-١١٠
٤٠-	٣٠	(٤٢,٥)	صفر	صفر
٤٥-	٣٤	٤٧,٥	+٥	+١٧٠
٥٠-٥٥	٦	٥٢,٥	+١٠	+٦٠
المجموع	١٠٠			-١٩٠ +٢٣٠
				+٤٠

ويكون حينئذ الوسط الحسابى من المعادلة (٣) فى الصورة:

$$\text{الوسط الحسابى} = \text{الوسط الفرضى} + \frac{\text{مجـ ح ك}}{\text{مجـ ك}}$$

$$= ٢٤,٥ + \frac{٤٠}{١٠٠} = ٤٢,٩ \text{ سنة}$$

وبالطبع سوف نحصل على نفس قيمة الوسط الحسابى لو إستخدمنا مراكز الفئات مباشرة والعلاقة (٢) فى إيجاد الوسط الحسابى. كما أننا سوف نحصل على نفس قيمة الوسط الحسابى لو إخترنا أى رقم آخر (سواء من بين مراكز الفئات السابقة أو من خارج الجدول) كوسط فرضى لتبسيط مراكز الفئات والحصول على انحرافات مراكز الفئات عن الوسط الفرضى.

وقد يرى البعض أنه يمكن أيضا تبسيط الانحرافات عن مراكـــز الفئـــات بقسمة جميع الانحرافات على مقدار ثابت فنحصل على الانحرافات المبســـطة لمراكز الفئات وسنرمز لها بالرمز ح. وبترجيح الانحراف المبسط لكل فئـــة بتكرار تلك الفئة وإستخدام العلاقة التالية لإيجاد الوسط الحسابى.

$$\text{الوسط الحسابى} = \text{الوسط الفرضى} + \text{المقدار الثابت} \times \frac{\text{مجـ ح}' \text{ ك}}{\text{مجـ ك}} \quad .....(٤)$$

مثال (٧)

يوضح الجدول التالى التوزيع التكرارى لعينة عشوائية مـــن ٨٠ متجـــرا موزعة طبقا لرأس المال لكل متجر بالألف جنيه.

رأس المال	٥٠-	٧٠-	٩٠-	١١٠-	١٣٠-	١٥٠-	المجموع
عدد المتاجر	١٠	٢٠	٢٥	١٥	٧	٣	٨٠

والمطلوب إيجاد الوسط الحسابى لـــرأس مـــال المتجـــر بإســـتخدام طريقـــة الانحرافات المبسطة.

الحـــل

بإيجاد مراكز الفئات أولا ثم اختيار ١٠٠ كوسط فرضى يتم حساب الانحراف ح عنه لكل مركز فئة نلاحظ أنه يمكن تبسيط جميع الانحرافات اذا قسمنا كـل انحراف على مقدار ثابت هو ط وبذلك نحصل على الانحرافات المبسـطة (حَ) وبترجيح الانحراف المبسط لكل فئة بتكرار الفئة وإستخدام العلاقة (٤) نحصل على الوسط الحسابى.

وبإعداد الجدول التالى.

رأس المال	ك	س	ح	ح/	ح/ك
٥٠-	١٠	٦٠	٤٠-	٢-	٢٠-
٧٠-	٢٠	٨٠	٢٠-	١-	٢٠-
٩٠-	٢٥	١٠٠	صفر	صفر	صفر
١١٠-	١٥	١٢٠	٢٠+	١+	١٥+
١٣٠-	٧	١٤٠	٤٠+	٢+	١٤+
١٥٠-	٣	١٦٠	٦٠+	٣+	٩+
المجموع	٨٠				٤٠- ٣٨+ ٢-

وباستخدام المعادلة (٤)

الوسط الحسابى = الوسط الفرضى + $\frac{\text{المقدار الثابت} \times \text{مجـ ح/ك}}{\text{مجـ ك}}$

$= ١٠٠ + \frac{٢٠ \times ٢-}{٨٠} = ٩٩,٥$ ألف جنيه

ومن البديهى أننا سنحصل على نفس قيمة الوسط الحسابى سواء استخدمنا مراكز الفئات مباشرة أو سواء استخدمنا طريقة الانحراف عن وسط فرضى أو سواء استخدمنا طريقة الانحراف المبسطة ذلك لأن الوسط الحسابى يخضع للمعالجة الجبرية ولذلك فيمكن تبسيط القيم بالصورة المناسبة ومعالجة تلك رياضيا باستخدام القانون المناسب للوسط الحسابى

ب - الوسيط:

يعتبر الوسيط المقياس التالى فى الأهمية للوسط الحسابى ويمكن أن يعرف بأنه قيمة المفردة الوسطى للقيم وذلك بشرط ترتيب القيم تصاعديا أو تنازليا.

وبناء على التعريف السابق للوسيط فإنه يتم إيجاد الوسيط بترتيب القيم تصاعديا أو تنازليا ثم بعد ذلك نعتبر أن الوسيط هو القيمة الوسطى تماما للقيم بعد إجراء هذا الترتيب وهذا يعنى أن عدد القيم التى تقل عن أو تساوى قيمة الوسيط مساو لعدد القيم التى تزيد عن أو تساوى قيمة الوسيط.

مثال (٨)

أوجد الوسيط من القيم التالية والتى توضح الوزن بالكيلو جرام لعينة عشوائية من سبعة طلاب من كلية التجارة.

٧٠٫٥، ٦٥٫٠٠ ، ٨٢٫٠٠ ، ٧١٫٥ ، ٧٥٫٥ ، ٨٤٫٥ ، ٧٨٫٠٠

الحـــل

نبدأ بترتيب القيم السابقة ترتيبا تصاعديا أو ترتيبا تنازليا:

الترتيب التصاعدى: ٦٥٫٠٠ ، ٧٠٫٥ ، ٧١٫٥ ، ٧٥٫٥ ، ٧٨٫٠٠ ، ٨٢٫٠٠ ، ٨٤٫٥.

الترتيب التنازلى : ٨٤٫٥ ، ٨٢٫٠٠ ، ٧٨٫٠٠ ، ٧٥٫٥ ، ٧١٫٥ ، ٧٠٫٥ ، ٦٥٫٠٠

وبالبحث عن القيمة الوسطى بين القيم المرتبة السابقة نجد أن هذه القيمة الوسطى تماما تساوى ٧٥٫٥ حيث هى القيمة ذات الرتبة الرابعة ويقل عنها ثلاثة قيم ويزيد عنها ثلاثة قيم.

مثال (٩)

أوجد الوسيط من القيم التالية والتى توضح الدخل بالجنيه لعينة عشـــوائية من ستة موظفين بجامعة عين شمس.

٢١٠،٨، ١٥٠،٦ ، ٢٩٠،٠٠ ، ٣١٨،٥ ، ١٨٨،٤ ، ٣٢٠،٦.

الحـــل

بترتيب القيم السابقة ترتيبا تصاعديا أو تنازليا.

الترتيب التصاعدى: ١٥٠،٦ ، ١٨٨،٤ ، ٢١٠،٨ ، ٢٩٠،٠٠ ، ٣١٨،٥ ، ٣٢٠،٦.

الترتيب التنازلـــى: ٣٢٠،٦ ، ٣١٨،٥ ، ٢٩٠،٠٠ ، ٢١٠،٨ ، ١٨٨،٤ ، ١٥٠،٦.

وبالبحث عن القيمة الوسطى تماما بين القيم المرتبة تصاعديا أو تنازليـــا فنلاحظ أن عدد القيم السابقة زوجيا (ستة قيم) وبالتالى فإن هذه القيمة الوسطى غير موجودة مباشرة (كما حدث فى المثال السابق حيث كان عدد القيـــم فـــى المثال السابق فرديا مما سمح بوجود القيمة الوسطى مباشرة بين القيم المرتبة) ولكنها محصورة بين القيمتين ٢١٠،٨ ، ٢٩٠،٠٠ ولذلك يمكـــن إعتبـــار أن الوسيط هو الوسط الحسابى للقيمتين الوسطيتين:

$$\therefore \text{الوسيط} = \frac{(٢١٠،٨ + ٢٩٠،٠٠)}{٢} = ٢٥٠،٤ \text{ جنيه}$$

كما أن رتبة الوسيط هنا بين الرتبة الثالثة والرتبة الرابعة.

وبصفة عامة فإن قيمة الوسيط تكون القيمة ذات الرتبة

$$\frac{(ن + ١)}{٢}$$

وذلك بفرض وجود ن من القيم المرتبة ترتيبا تصاعديا أو تنازليا فإذا كانت ن فردية فإن قيمة الوسيط سنجدها مباشرة بين القيم المرتبة أما اذا كانت ن زوجية فإن قيمة ا لوسيط ستكون هى الوسط الحسابى للقيمتين الوسطيتين وذلك بشرط ترتيب القيم أولا بصورة تصاعدية أو تنازلية، ففى المثال (٨) كانت رتبة الوسيط هى

$$\frac{(ن + ١)}{٢} = \frac{(٧ + ١)}{٢} = ٤$$

أى أن قيمة الوسيط هى قيمة المفردة الرابعة فى الترتيب للمفردات المرتبة ترتيبا تصاعديا أو تنازليا بينما فى المثال (٩) كانت رتبة الوسيط هى

$$\frac{(ن + ١)}{٢} = \frac{(٦ + ١)}{٢} = ٣,٥$$

ومن المنطقى أن هذه الرتبة تقع بين الرتبتين ٣ ، ٤ للقيم المرتبة تصاعديا أو تنازليا.

ونلاحظ من مثال (٨)، مثال (٩) أن قيمة الوسيط هى قيمة المفردة الوسطى تماما بحيث يوجد نصف القيم قبل قيمة الوسيط والنصف الآخر للقيم يوجد بعد الوسيط وذلك للقيم المرتبة ترتيبا تصاعديا أو ترتيبا تنازليا وقد كانت القيم لمتغير متصل فى كلا المثالين، ولكن عند إيجاد الوسيط لمتغير منفصل حيث من الممكن أن تتكرر أى قيمة لعدة مرات فإننا قد لانجد أن نصف القيم المرتبة تصاعديا أو تنازليا تكون قبل قيمة الوسيط والنصف الآخر من القيم بعد قيمة الوسيط

مثال (١٠)

أوجد الوسيط من القيم التالية والتى توضح عـدد أفـراد الأسـرة لعينـة عشوائية من سبعة طلاب من كلية التجارة: ٩ ، ٥ ، ٦ ، ٤ ، ٥ ، ٤ ، ٥.

الحــل

بترتيب القيم السابقة ترتيبا تصاعديا أو تنازليا نجد أن:

الترتيب التصاعدى: ٤ ، ٤ ، ٥ ، ٥ ، ٥ ، ٦ ، ٩.

الترتيب التنازلـى: ٩ ، ٦ ، ٥ ، ٥ ، ٥ ، ٤ ، ٤.

وتكون رتبة الوسيط هى

$$\frac{(٧ + ١)}{٢} = ٤$$

أى أن قيمة الوسيط هى قيمة المفردة الرابعة فى الترتيب وهى هنا تسـاوى ٥ ولكن نظرا لتكرار هذه القيمة فإننا نجد أن نصف القيم المرتبـة تقـع قبلـها والنصف الأخر يقع بعدها.

إيجاد الوسيط من التوزيع التكرارى:

يمكن إيجاد الوسيط من التوزيع التكرارى بنفس الفلسفة السابقة للبيانـات غير المبوبة فيجب أولا ترتيب القيم ترتيبا تصاعديا أو ترتيبا تنازليا ثم إيجـاد رتبة الوسيط باعتباره رتبة القيمة الوسطى تماما والتى يقل عنها نصف القيـم ويزيد عنها نصف القيم. ولكن الجدول التكرارى لايعبر مباشرة عـن القيـم الحقيقية للظاهرة (كما أوضحنا فى تبويب القيم فى جدول تكـرارى) ولكننـا سنفترض أن القيم التى تقع داخل كل فئة فى الجدول التكرارى تكون موزعـة

توزيعا منتظما داخل الفئة، ومن ثم يمكننا إيجاد قيمة الوسيط باعتبارە قيمة المفردة الوسطى تماما من القيم المرتبة ترتيبا تصاعديا أو تنازليا وذلك بالبحث بطريقة النسبة والتناسب عن هذه القيمة الوسطى داخل الفئة الوسيطية والتي يقع فيها الوسيط.

ويمكننا من خلال هذا الاسلوب الحصول على تقدير جيد لقيمة الوسيط وتزداد جودة هذا التقدير كلما صغر طول الفئة الوسيطية. وبذلك يمكن تلخيص خطوات إيجاد الوسيط من التوزيع التكرارى كما يلى:-

١) ترتيب القيم تصاعديا او تنازليا وذلك بتكوين التوزيع التكرارى المتجمع الصاعد أو التوزيع التكرارى المتجمع الهابط.

٢) إيجاد رتبة الوسيط بإعتبار أن الوسيط هو القيمة الوسطى تماما وبذلك تكون رتبة الوسيط تساوى نصف مجموع التكرارات أى أن

$$\text{رتبة الوسيط} = \frac{\text{مجـ ك}}{٢} \qquad \ldots\ldots\ldots\ldots(٥)$$

٣) نبحث عن رتبة الوسيط بين التكرار المتجمع وذلك لتحديد الفئة الوسيطية وبذلك يكون قد تحدد الحد الأدنى والحد الأعلى لقيمة الوسيط.

٤) بافتراض أن القيم داخل الفئة الوسيطية موزعة توزيعا منتظما وباستخدام أسلوب النسبة والتناسب يمكن تحديد الجزء الذى يجب إضافته الى بداية الفئة الوسيطية لنحصل على قيمة الوسيط وحيث هذا الجزء يساوى:

$$\frac{\text{طول الفئة الوسيطية} \times \text{الفرق بين رتبة الوسيط والتكرار المتجمع السابق له}}{\text{تكرار الفئة الوسيطية}}$$

$$\text{الوسيط} = \text{بداية الفئة الوسيطية} + \frac{\text{طول الفئة الوسيطية} \times \text{الفرق بين رتبة الوسيط والتكرار المتجمع السابق له}}{\text{تكرار الفئة الوسيطية}} \quad ....(٦)$$

مثال (١١)

يوضح الجدول التالى التوزيع التكرارى للعمر لعينة عشوائية مــــن ١٠٠ موظف من جامعة عين شمس:

فئات العمر	٣٠-	٣٥-	٤٠-	٤٥-	٥٠-	٥٥-٦٠
عدد الموظفين	٨	١٢	٣٦	٢٤	١٥	٥

والمطلوب : إيجاد الوسيط

الحـــــل

١) بتكوين التوزيع التكرارى المتجمع الصاعد نجد أن

أقل من حدود الفئات	التكرار المتجمع الصاعد
أقل من ٣٠	صفر
أقل من ٣٥	٨
أقل من ٤٠	٢٠
أقل من ٤٥	٥٦
أقل من ٥٠	٨٠
أقل من ٥٥	٩٥
أقل من ٦٠	١٠٠

٢) تحديد رتبة الوسيط حيث :

$$\text{رتبة الوسيط} = \frac{\text{مجـ ك}}{٢} = \frac{١٠٠}{٢} = ٥٠$$

وبالتالى يكون الوسيط هو قيمة المفردة الخمسين فى الجدول.

٣) بالبحث عن رتبة الوسيط نجد أنها محصورة بين التكرار المتجمع الصاعد ٢٠، ٥٦ وبذلك تكون قيمة الوسيط محصورة بين ٤٠، ٤٥ وتكون قيمة الوسيط هي:

$$\text{قيمة الوسيط} = ٤٠ + \frac{٥\,(٥٠ - ٢٠)}{٣٦}$$

$$= ٤٠ + ٤{,}٢ = ٤٤{,}٢ \text{ سنة}$$

ونلاحظ أن تكرار الفئة الوسيطية = الفرق بين التكرار المتجمع اللاحق لرتبة الوسيط والتكرار المتجمع السابق لرتبة الوسيط.

واذا اعتبرنا أن:

ف = بداية الفئة الوسيطية

ط = طول الفئة الوسيطية

$ك_٢$ = رتبة الوسيط

$ك_١$ = التكرار المتجمع السابق لرتبة الوسيط (ك. م. ص السابق)

$ك_٣$ = التكرار المتجمع اللاحق لرتبة الوسيط (ك.م.ص اللاحق)

يمكن القول أن:

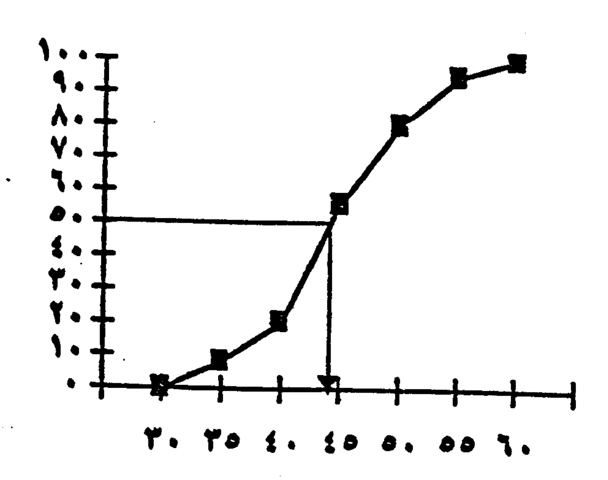
$$\text{قيمة الوسيط} = \text{ف} + \frac{\text{ط}\,(ك_٢ - ك_١)}{(ك_٣ - ك_١)} \quad \text{.....................(٧)}$$

ومن الجدول السابق نجد أن:

$$\text{الوسيط} = ٤٠ + \frac{٥\,(٥٠ - ٢٠)}{٥٦ - ٢٠}$$

$$= ٤٠ + ٤{,}٢ = ٤٤{,}٢ \text{ سنة}$$

ويمكن إيجاد الوسيط من خلال الرسم وذلك برسم المنحنى المتجمع الصاعد، ثم نبحث عن رتبة الوسيط على المحور الرأسى بين التكرار المتجمع ونقوم برسم خط موازى للمحور الأفقى يبدأ من رتبة الوسيط حتى يلتقى مع المنحنى المتجمع الصاعد ومنها نسقط عمودا رأسيا حيث تتحدد قيمة الوسيط بنقطة إلتقاء العمود الرأسى مع المحور الأفقى ويتضح ذلك من الشكل التالى:



شكل (١)

ويمكن استخدام التوزيع التكرارى المتجمع الهابط (الترتيب التنازلى) لإيجاد قيمة الوسيط كما يلى:

حدود الفئات فأكثر	التكرار المتجمع الهابط
٣٠ فأكثر	١٠٠
٣٥ فأكثر	٩٢
٤٠ فأكثر	٨٠
٤٥ فأكثر	٤٤
٥٠ فأكثر	٢٠
٥٥ فأكثر	٥
٦٠ فأكثر	صفر

وحيث أن رتبة الوسيط =

= مجـ ك / ٢ = ١٠٠ / ٢ = ٥٠

وبالبحث عن رتبة الوسيط بين التكرار المتجمع الهابط نلاحظ أن:

التكرار المتجمع السابق لرتبة الوسيط = ٨٠ = ك١

التكرار المتجمع اللاحق لرتبة الوسيط = ٤٤ = ك٢

وتكون قيمة الوسيط محصورة بين ٤٠، ٤٥.

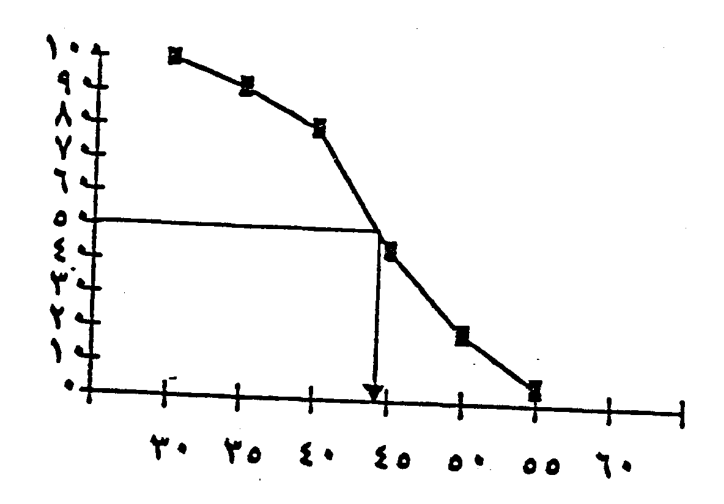
وباستخدام القانون التالى لايجاد الوسيط طالما أن ذلك يتم من التكرار المتجمع الهابط.

الوسيط = بداية الفئة الوسيطية + (طول فئة الوسيط × الفرق بين رتبة الوسيط والتكرار المتجمع السابق) / تكرار الفئة الوسيطية .....(٨)

الوسيط = ف + ط × (ك٢ − ك١) / (ك٢ − ك١) ....(٩)

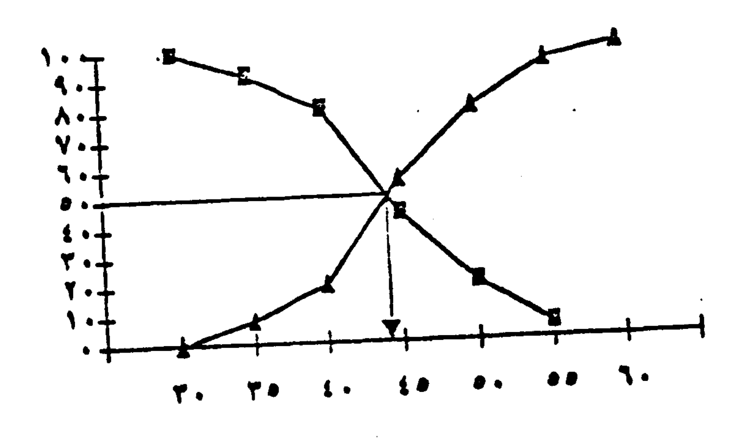
= ٤٠ + ٥ × (٥٠ − ٨٠) / (٤٤ − ٨٠) = ٤٤,٢ سنة

ويمكن إيجاد الوسيط بالرسم من التوزيع التكرارى المتجمع الهابط باستخدام المنحنى المتجمع الهابط وذلك بنفس الطريقة التى أوضحناها فى المنحنى المتجمع الصاعد - ويتضح ذلك من الشكل التالى:



شكل (٢)

ويمكن إيجاد الوسيط من خلال رسم المنحنى المتجمع الصاعد والمنحنى المتجمع الهابط فى شكل واحد حيث نسقط عمودا على المحور الأفقى من نقطة لقاء المنحنى المتجمع الصاعد والمنحنى المتجمع الهابط لتحديد قيمة الوسيط ويتضح ذلك من الشكل التالى.



شكل (٣)

ويمكن إيجاد الوسيط من الجداول التكرارية غير المنتظمة الفئات او الجداول التكرارية المفتوحة من أحد طرفيها أو من كليهما وتعتبر هذه أحد السمات المميزة للوسيط حيث أن الوسيط هو المتوسط ومقياس النزعة المركزية الذى قد يعبر عن الظاهرة بدقة عندما تكون قيم الظاهرة قد تم تبويبها فى جدول تكرارى مفتوح.

مثال (١٢)

يوضح التوزيع التكرارى التالى درجات عينة عشوائية من ٢٠٠ طالب فى كلية التجارة فى مادة الاحصاء.

الدرجة	٢٥-	٤٠-	٦٠-	٧٠-	٨٠-	٩٠-١٠٠
عدد الطلاب	١٥	٤٩	٧٥	٣١	٢٥	٥

أوجد الوسيط بالحساب وبالرسم باستخدام التوزيع التكرارى المتجمع الصاعد

الحــــل

أقل من حدود الفئات	التكرار المتجمع الصاعد
أقل من ٢٥	صفر
أقل من ٤٠	١٥
أقل من ٦٠	٦٤
أقل من ٧٠	١٣٩
أقل من ٨٠	١٧٠
أقل من ٩٠	١٩٥
أقل من ١٠٠	٢٠٠

$$\text{رتبة الوسيط} = \frac{\text{مجـ ك}}{٢} = \frac{٢٠٠}{٢} = ١٠٠$$

وبالبحث عن رتبة الوسيط فى التكرار المتجمع الصــاعد نجـد أنــها محصورة بين ٦٤، ١٣٩ ومن ثم فإن قيمة الوسيط تقع بين ٦٠، ٧٠ درجــة ويكون:

رتبة الوسيــــط = $ك_٢$ = ١٠٠

ك. م. ص السابق = $ك_١$ = ٦٤

ك. م. ص اللاحق = $ك_٣$ = ١٣٩

ويكون تكرار الفئة الوسيطية = ك = $ك_٣$ - $ك_١$ = ١٣٩ - ٦٤ = ٧٥

بداية الفئة الوسيطية = ف = ٦٠

طول الفئة الوسيطية = ط = ٧٠ - ٦٠ = ١٠

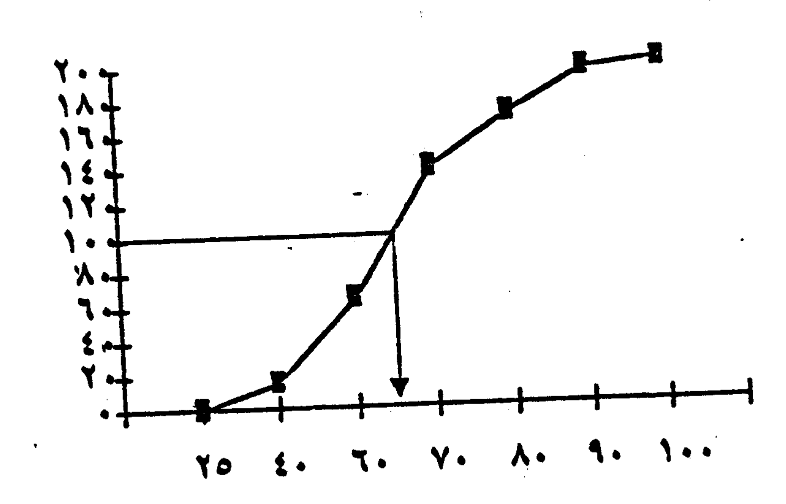
وباستخدام المعادلة (٧)

$$\text{الوسيط} = \text{ف} + \frac{\text{ط}\,(\text{ك}_{2} - \text{ك}_{1})}{\text{ك}_{3} - \text{ك}_{1}}$$

$$= ٦٠ + \frac{١٠\,(١٠٠ - ٦٤)}{١٣٩ - ٦٤}$$

$= ٦٠ + ٤,٨ = ٦٤,٨$ درجة

وبرسم المنحنى المتجمع الصاعد يمكن إيجاد قيمة الوسيط من الرسم كما يتضح من الشكل التالى:



شكل (٤)

مثال (١٣)

يوضح التوزيع التالى بيانات عينة عشوائية من ٣٠٠ طالب موزعة حسب فئات الذكاء

فئات الذكاء	أقل من ٦٠	٦٠-	٧٠-	٧٥-	٨٥-	١٠٠-١٢٠
عدد الطلاب	٢٢	٥٨	١٠٠	٦٠	٤٥	١٥

والمطلوب:

إيجاد الوسيط بالحساب باستخدام التوزيع التكرارى المتجمع الصاعد

الحـــل

أقل من حدود الفئات	التكرار المتجمع الصاعد
أقل من ٦٠	٢٢
أقل من ٧٠	٨٠
أقل من ٧٥	١٨٠
أقل من ٨٥	٢٤٠
أقل من ١٠٠	٢٨٥
أقل من ١٢٠	٣٠٠

رتبة الوسيط= $\frac{\text{مجـ ك}}{٢}$ = $\frac{٣٠٠}{٢}$ = ١٥٠

وبالبحث فى التكرار المتجمع الصاعد نجد أن رتبة الوسيط تقع بيـــن ٨٠، ١٨٠ ومن ثم تكون قيمة الوسيط محصورة بين ٧٠ ، ٧٥ وتكون:

رتبة الوسيـــط = ك$_{٢}$ = ١٥٠

ك.م.ص. السابق = ك$_{١}$ = ٨٠

ك.م.ص. اللاحق = ك$_{٣}$ = ١٨٠

تكرار الفئة الوسيطية = ك = ك$_{٣}$ - ك$_{١}$ = ١٨٠ - ٨٠ = ١٠٠.

بداية الفئة الوسيطية = ف = ٧٠

طول الفئة الوسيطية = ط = ٧٥ − ٧٠ = ٥

$$\therefore \text{الوسيط} = \text{ف}_{٠} + \frac{\text{ط} (\text{ك}_٢ - \text{ك}_١)}{\text{ك}_٣ - \text{ك}_١}$$

$$= ٧٠ + \frac{٥ (١٥٠ - ٨٠)}{١٨٠ - ٨٠}$$

= ٧٠ + ٣,٥ = ٣٧,٥ درجة

ويحدث فى بعض الأحيان عند إيجاد قيمة الوسيط من التوزيع التكرارى المتجمع الصاعد أو التوزيع التكرارى المتجمع الهابط أن نلاحظ مباشرة أن رتبة الوسيط التى نبحث عنها بين التكرارات المتجمعة موجودة مباشرة وبالضبط لأحد التكرارات المتجمعة وهذا يعنى أن قيمة الوسيط التى نبحث عنها توجد مباشرة فى التوزيع التكرارى المتجمع حيث تكون قيمة الوسيط هى الحد الذى يقابل التكرار المتجمع الذى هو رتبة الوسيط.

مثال (١٤)

أوجد الوسيط من التوزيع التكرارى التالى الذى يوضح توزيع عينة عشوائية من ١٢٠ طفل حسب فئات الوزن المختلفة.

فئات الوزن	١٠−	١٦−	٢٢−	٢٨−	٩٠−١٠٠
عدد الأطفال	٥	١٥	٤٠	٤٧	١٣

الحــــــل

أقل من حدود الفئات	ك . م . ص
أقل من ١٠	صفر
أقل من ١٦	٥
أقل من ٢٢	٢٠
أقل من ٢٨	٦٠
أقل من ٣٤	١٠٧
أقل من ٤٠	١٢٠

رتبة الوسيط= $\frac{\text{مجـ ك}}{٢} = \frac{١٢٠}{٢} = ٦٠$

وبالبحث عن رتبة الوسيط بين التكرارات المتجمعة نجد أن هذه الرتبة هي أحد التكرارات المتجمعة فى التوزيع المتجمع الصاعد وهذا يعني أن قيمة الوسيط مباشرة هي ٢٨ كيلو جرام.

ونود أن نؤكد أن ايجاد قيمة الوسيط مباشرة من قيم الظاهرة الخام غير المبوبة يعطى القيمة الدقيقة للوسيط فى حين أن قيمة الوسيط التي نحصل عليها من التوزيع التكرارى للظاهرة تكون قيمة تقريبية للوسيط.

خصائص الوسيط:

١) تعتبر الوسيط من أهم المتوسطات ويمكن القول أنه يلى الوسط الحسابى فى الأهمية فى معظم الظواهر بل وقد نعتبره أفضل من الوسط الحسابى كمتوسط للظاهرة فى بعض الأحيان وذلك حينما توجد بعض القيم المتطرفة بين مختلف قيم الظاهرة وذلك لأن الوسيط لايتأثر بالقيم المتطرفة، كما أنه يمكن استنتاج قيمةالوسيط بالرسم، وبالاضافة الى ذلك فإن الوسيط يتميز بإمكانية ايجاد قيمته من الجداول التكرارية المفتوحة.

٢) يعتبر الوسيط مقياسا لمنازل القيم.

٣) لايخضع للمناولة الجبرية بعكس الوسط الحسابى.

٤) مجموع الانحرافات المطلقة للقيم عن الوسيط أقل من مجموع الانحرافات المطلقة للقيم عن الوسط الحسابى أو عن أى قيمة أخرى.

ويمكن توضيح هذه الخاصية من خلال المثال التالى:

مثال (١٥)

اذا أعطيت القيم التالية:

١٠، ١٥، ٨، ١٢، ٥، ١٤، ١٣

أوجد الوسط الحسابى والوسيط ثم أوجد مجموع الانحرافات المطلقة للقيم عن كل منهما.

الحل

$$\bar{س} = \frac{مجـ س}{ن} = \frac{٧٧}{٧} = ١١$$

وبترتيب القيم تصاعديا:

٥ ٨ ١٠ ١٢ ١٣ ١٤ ١٥

ومن ثم فإن الوسيط = ١٢

ونجد أن مجموع الانحرافات المطلقة للقيم عن الوسط الحسابى

= ٦ + ٣ + ١ + ١ + ٢ + ٣ + ٤ = ٢٠

كما أن مجموع الانحرافات المطلقة للقيم عن الوسيط

= ٧ + ٤ + ٢ + صفر + ١ + ٢ + ٣ = ١٩

واذا اخترنا قيمة أكبر من الوسيط ولنفترض أنها ١٤ وبايجاد مجموع الانحرافات المطلقة للقيم عن ١٤ نجده

= ٩ + ٦ + ٤ + ٢ + ١ + صفر + ١ = ٢٣

ويتضح من ذلك أن مجموع الانحرافات المطلقة للقيم عن الوسيط أقـــل مـــن مجموع الانحرافات المطلقة للقيم عن الوسط الحسابى أو عن أى قيمة أخرى.

جـ - المنــوال :

يمكن أن نعرف المنوال بأنه القيمة الأكثر شيوعا وتكرارا. ويعنـــى هــذا التعريف للمنوال ضرورة أن تكون هناك قيمة شائعة بمعنى أنها قيمة تكــورت أكثر من غيرها من القيم. فإذا اعتبرنا القيم التالية:

٨، ٢ ، ١٠، ١٥، ٧، ٩، ٣٤

فإننا نقول أنه يمكن إيجاد الوسط الحسابى أو الوسيط لهذه القيم ولكن لايوجــد منوال لهذه القيم لأنه لم تتكرر أحد القيم ولاتوجد قيمة شائعة بين القيم.

ولنعتبر الجدول التالى الذى يوضح توزيع عينة من ٢٠٠ طالب بكلية التجــارة حسب عدد أفراد الأسرة:

عدد أفراد الأسرة	٢	٣	٤	٥	٦	٧
عدد الطلاب	١٠	٢٥	١١٠	٤٥	٧	٣

هنا يمكن القول أن المنوال يساوى ٤ أفراد وذلك المقابل لأكبر تكرار (١١٠)

ويمكننا القول أن المنوال يكون أحد قيم الظاهرة وذلك طبقا للتعريف فــهو القيمة التى نجد أنها قد تكررت بصورة أكثر من غيرها من القيم ولـــهذا قــد يحدث ألا نجد منوالا للظاهرة وذلك عندما لاتتكرر أحد القيم وهنا لانســتطيع إيجاد المنوال ويمكن حينئذ إيجاد الوسط الحسابى أو الوسيط كمقياس للنزعــة المركزية للظاهرة، ولهذا لايعتبر المنوال متوسطا جيدا الا اذا كانت هناك قيمة شائعة ومتكررة فى الظاهرة بصورة واضحة.

ويمكن إيجاد المنوال من التوزيع التكرارى بعدة طرق وجميع هذه الطرق تعطى قيما تقريبية للمنوال، فعندما تتوافر لدينا بيانات عن متغير متصل فــإن القيم المختلفة للمتغير المتصل لاتسمح عادة بوجود قيمــة متكــررة نعتبرهــا المنوال، كما أن قيم المتغير يتم تبويبها فى توزيع تكــرارى يتضمــن فئــات المتغير والتكرارات المقابلة لكل فئة ومن ثم سوف نحدد الفئة التى تقابل أكــبر تكرار بإعتبارها الفئة المنوالية أى التى تحتوى على القيمة الشائعة للمنــوال، فإذا ماأعدنا تبويب قيم الظاهرة فى توزيع تكرارى تختلف فئاته عن التوزيــع الأول فقد نحصل على فئة منوالية جديدة وهكذا.

ولهذا نرى أننا نحصل على قيمة تقريبية للمنوال عند إيجاده من التوزيــع التكرارى للمتغير المتصل.

وهناك أكثر من طريقة تقريبية لإيجاد المنوال:

١- طريقة مركز الفئة المنوالية:

تعتمد هذه الطريقة فى تقدير المنوال على مركز الفئة المنوالية أى علــى مركز الفئة المقابلة لأكبر تكرار . ومن البديهى أن يكون الجدول التكــرارى متساوى ومنتظم الفئات عند تحديد الفئة المنوالية المقابلة لأكــبر تكــرار وإلا يجب تحديد الفئة المنوالية بناء على أكبر تكرار معدل وذلك عندمــا يكــون الجدول غير منتظم الفئات.

مثال (١٦)

يوضح التوزيع التكرارى التالى توزيع عينة من ٨٠ طالب حسب العمر:

العمر	١٤-	١٦-	١٨-	٢٠-	٢٢-٢٤
عدد الطلاب	٥	٢٥	٣٠	١٥	٥

أوجد المنوال للعمر

الحـــل

حيث أن أكبر تكرار = ٣٠

فإن المنوال يقع فى الفئة المنوالية المقابلة لأكبر تكرار أى أن قيمة المنوال تقع بين ١٨ الى أقل من ٢٠ سنة وبإعتبار أن:

المنوال = مركز الفئة المنوالية

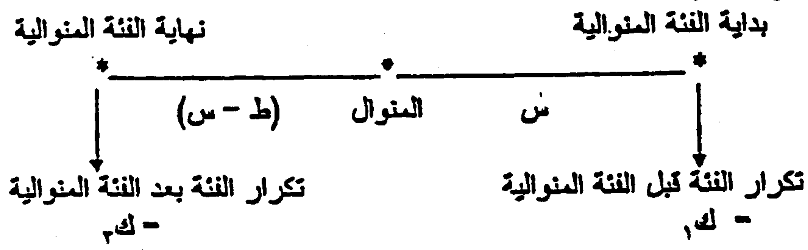
$$= \frac{\text{(بداية الفئة المنوالية + نهاية الفئة المنوالية)}}{٢}$$

$$= \frac{(١٨ + ٢٠)}{٢} = ١٩ \text{ سنة}$$

٢- طريقة الرافعة:

تفترض هذه الطريقة أن هناك قوتان تتفاعلان فى الفئة المنوالية ويستقر المنوال عند نقطة الاتزان الناتجة بعد تفاعل تلك القوتين. أول هاتان القوتان هو التكرار السابق لأكبر تكرار (أى تكرار الفئة التى قبل الفئة المنوالية) وتعمل هذه القوة على شد المنوال الى بداية الفئة المنوالية بينما القوة الثانية هى التكرار التالى (اللاحق) لأكبر تكرار (أى تكرار الفئة التى بعد الفئة المنوالية) وتعمل هذه القوة على شد المنوال الى نهاية الفئة المنوالية ومن ثم تتحدد قيمة

المنوال داخل الفئة المنوالية نتيجة لتفاعل تلك القوتان وذلك من خلال قـــانون الرافعة حيث:



وباعتبار أن:

القوة × ذراعها = المقاومة × ذراعها

وباعتبار أن المسافة من بداية الفئة المنوالية وحتى نقطة الاتـــزان التـى سيستقر عندها المنوال = س

فـــإن:

تكرار الفئة قبل الفئة المنوالية × س = تكرار الفئة بعد المنوالية × (ط - س)

أى أن:

ك١ × س = ك٢ (ط - س)

اذا س (ك١ + ك٢) = ك٢ ط

$$\text{ومنها س} = \frac{\text{ك}_2\ \text{ط}}{(\text{ك}_1 - \text{ك}_2)}$$

وتكون قيمة المنوال = بداية الفئة المنوالية + س

$$= \text{بداية الفئة المنوالية} + \frac{\text{ك}_2\ \text{ط}}{(\text{ك}_1 + \text{ك}_2)} \quad \text{.................(١٠)}$$

مثال (١٧)

أوجد قيمة المنوال للعمر بطريقة الرافعة من بيانات المثال (١٦)

حيث أن أكبر تكرار = ٣٠

فإن الفئة المنوالية من ١٨ الى أقل من ٢٠ ومن ثم فإن طول الفئة المنوالية = ط = ٢

ويكون التكرار السابق لتكرار الفئة المنوالية = $ك_{١}$ = ٢٥

ويكون التكرار اللاحق لتكرار الفئة المنوالية = $ك_{٣}$ = ١٥

ومن المعادلة (١٠)

$$\text{المنوال} = \text{بداية الفئة المنوالية} + \frac{ك_{٣}\ ط}{(ك_{١} + ك_{٣})}$$

$$= ١٨ + \frac{١٥ \times ٢}{٢٥ + ١٥} = ١٨{,}٧٥ \text{ سنة}$$

ويجب أن نلاحظ أن تكرار الفئة المنوالية (أكبر تكرار) وسنرمز له بالرمز $ك_{٢}$ لم يكن له أى دور فى تحديد قيمة المنوال داخل الفئة المنوالية ولكنه فقط تم من خلاله تحديد الفئة المنوالية.

٣- طريقة فروق التكرارات:

وتتفق هذه الطريقة مع الطرق السابقة فى أن الفئة المنوالية هى الفئة التى تقابل أكبر تكرار طالما أن الجدول التكرارى منتظم ومتساوى الفئات أما لو كان الجدول التكرارى غير منتظم الفئات فإن الفئة المنوالية هى الفئة التى تقابل أكبر تكرار معدل ولكنها تختلف عن الطرق السابقة فى أنها تعتمد على الفرق بين أكبر تكرار والتكرار السابق له وعلى الفرق بين أكبر تكرار

والتكرار اللاحق له وذلك عند تحديد نقطة الاتزان التى استقر عندها المنوال فى الفئة المنوالية.

ويطلق عادة على الفرق بين أكبر تكرار والتكرار السابق له الفرق الأول وسنرمز له بالرمز $ف_١$ حيث أن $ف_١ = ك_٢ - ك_١$ كما يطلق على الفرق بين أكبر تكرار والتكرار اللاحق له الفرق الثانى وسنرمز له $ف_٢$ حيث $ف_٢ = ك_٢ - ك_٣$

ويكون المنوال حينئذ :

$$\text{المنوال} = \text{بداية الفئة المنوالية} + \frac{ط\ ف_١}{ف_١ + ف_٢} \quad ....(١١)$$

مثال (١٨)

أوجد منوال العمر بطريق فروق التكرارات من التوزيع التكرارى فى مثال (١٦)

الحـــل

نلاحظ أن الفئة المنوالية من ١٨ الى ٢٠ سنة

ويكون:

الفرق الأول = $ف_١ = ك_٢ - ك_١ = ٣٠ - ٢٥ = ٥$

الفرق الثانى = $ف_٢ = ك_٢ - ك_٣ = ٣٠ - ١٥ = ١٥$

ومن المعادلة (١١)

$$\therefore \text{المنوال} = \text{بداية الفئة المنوالية} + \frac{ط\ ف_١}{ف_١ + ف_٢}$$

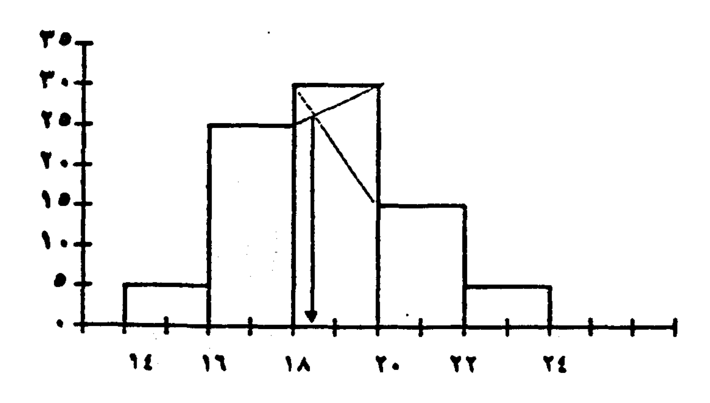
$$= ١٨ + \frac{٢ \times ٥}{٥ + ١٥} = ١٨,٥ \text{ سنة}$$

ونذكر القارىء بأننا أوضحنا السمة التقريبية للمنوال ومن ثم فإنه من الطبيعى أنه قد تختلف قيمة المنوال فى الطرق المختلفة.

٤- تقدير المنوال بالرسم:

يمكن تقدير قيمة المنوال بالرسم من خلال رسم المدرج التكرارى وتحديد المنوال بإسقاط عمود من نقطة تقاطع الخط المستقيم الواصل بين نقطة إلتقاء أكبر تكرار والتكرار التالى بقيمة أكبر تكرار وبين الخط المستقيم الواصل بين نقطة إلتقاء أكبر تكرار والتكرار السابق له بقيمة أكبر تكرار. ويتضح ذلك من الرسم التالى والذى يوضح إيجاد المنوال بالرسم من التوزيع التكرارى فى مثال (١٦).

وجدير بالذكر أن قيمة المنوال بالرسم بهذه الطريقة تتفق دائما مع قيمة المنوال بطريقة فروق التكرارات.



شكل (٥)

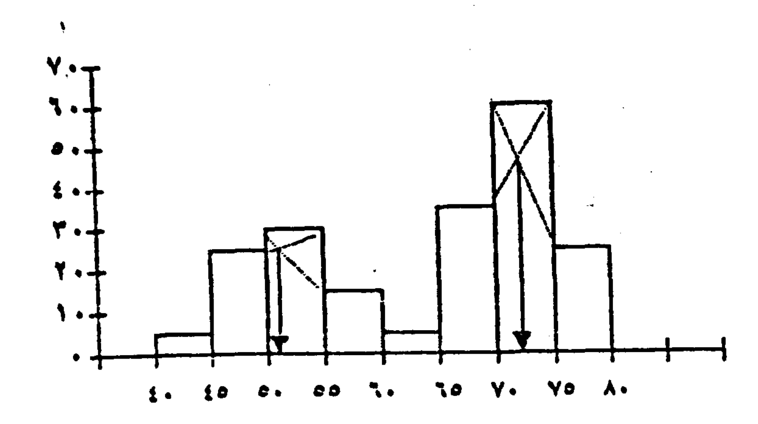
ويحدث فى بعض الأحيان أن يكون مجتمع الظاهرة يتكون فى حقيقة الأمر من مجتمعين مختلفين فى طبيعة مفردات كل منهم، فقد يكون مجتمع البحث مكون من ذكور وإناث أو مثلا من الريف والحضر. هنا من الممكن أن يوجد أكثر من منوال للظاهرة بحيث يكون هناك منوالا يعبر عن النزعة المركزية لمفردات كل مجتمع.

وبرسم المدرج التكرارى الذى يعبر عن التوزيع التكرارى نلاحظ أن هناك منوالان للظاهرة.

مثال(١٩)

يوضح الجدول التالى توزيع عينة من ٢٠٠ طالب وطالبة بكلية التجارة حسب فئات الوزن المختلفة – والمطلوب رسم المدرج التكرارى وإيجاد المنوال.

فئات الوزن	٤٠-	٤٥-	٥٠-	٥٥-	٦٠-	٦٥-	٧٠-	٨٠-٧٥
العدد	٥	٢٥	٣٠	١٥	٥	٣٥	٦٠	٢٥



شكل (٦)

يتضح من المدرج التكرارى أنه يوجد منوال للوزن فى الفئة من ٥٠ الى ٥٥ كيلو جرام بينما يوجد منوال آخر فى الفئة من ٧٠ الى ٧٥ كيلو جرام. ويمكن إيجاد كل منهما بأحد الطرق السابقة، فإذا استخدمنا طريقة الرافعة على سبيل المثال فإن:

$$\text{المنوال} = \text{بداية الفئة المنوالية} + \frac{\text{ط ك}_2}{\text{ك}_1 + \text{ك}_2}$$

$$\text{المنوال الأول} = ٥٠ + \frac{٥ \times ١٥}{٢٥ + ١٥} = ٥١,٩ \text{ كيلو جرام}$$

$$\text{المنوال الثانى} = ٧٠ + \frac{٢٥ \times ٥}{٣٥ + ٢٥} = ٧٢,١ \text{ كيلو جرام}$$

ويمكن إيجاد المنوال من التوزيع التكرارى غير المتساوى الفئات بنفس الطرق التى سبق عرضها إلا إننا يجب أن نستخدم التكرار المعدل عند إيجاد المنوال، حيث أن التكرار المعدل يعبر عن عدد التكرارات فى وحدة طول الفئة.

فحينما كانت الفئات متساوية تم إستخدام التكرارات مباشرة وبحثنا عن أكبر تكرار حتى يمكن تحديد الفئة المنوالية ولكن حينما تختلف أطول فئات التوزيع التكرارى فإن أكبر تكرار حينئذ قد يشير وقد لا يشير الى الفئة المنوالية ولكن عند إيجاد عدد التكرارات فى وحدة طول الفئة من خلال إيجاد التكرار المعدل، فإن أكبر تكرار معدل حينئذ سوف يشير الى الفئة المنوالية.

مثال (٢٠)

أوجد المنوال من التوزيع التكرارى التالى الذى يوضح توزيع عينة من ٥٠٠ طالب على فئات الطول المختلفة:

فئات الطول	-١٥٠	-١٥٥	-١٦٠	-١٦٥	-١٧٠	-١٨٠	١٩٠-٢٠٠
عدد الطلاب	٤٠	٥٠	٨٠	١١٠	١٣٠	٧٠	٢٠

الحـــل

نلاحظ أن أكبر عدد للطلاب هو ١٣٠ مما قد يوحى بأن الفئة المنوالية من ١٧٠ الى ١٨٠ ولكن هذا بالطبع غير سليم نظرا لأن التوزيع التكرارى غـير متساوى الفئات ويجب أن نوجد التكرار المعدل الذى يعبر عن عدد التكرارات (عدد الطلاب) فى وحدة طول الفئة وحيث نفترض أن التكـــرارات موزعـة بصورة منتظمة داخل كل فئة :

فئات الطول	عدد الطلاب	طول الفئة	التكرار المعدل
-١٥٠	٤٠	٥	٨
-١٥٥	٥٠	٥	١٠
-١٦٠	٨٠	٥	١٦
-١٦٥	١١٠	٥	٢٢
-١٧٠	١٣٠	١٠	١٣
-١٨٠	٧٠	١٠	٧
١٩٠-١٩٨	٢٠	٨	٢,٥
المجموع	٢٠٠		

ونجد أن أكبر تكرار معدل هو ٢٢ ومن ثم فإن الفئة المنوالية من ١٦٥ الـــى ١٧٠ سنتيمتر ويكون:

ك١ = ١٦

ك٢ = ١٣

$$\text{المنوال بطريقة الرافعة} = \text{بداية فئة المنوال} + \frac{\text{ط ك}_٢}{\text{ك}_١ + \text{ك}_٢}$$

$$= ١٦٥ + \frac{٥ \times ١٣}{١٦ + ١٣} = ١٦٧,٢ \text{ سنتيمتر}$$

ويمكن إيجاد المنوال بطريقة فروق التكرارات حيث:

ك٣ = ٢٢

الفرق الأول = ف١ = ك٣ ــ ك١ = ٢٢ ــ ١٦ = ٦

الفرق الثانى = ف٢ = ك٣ ــ ك٢ = ٢٢ ــ ١٣ = ٩

$$\text{المنوال} = \text{بداية فئة المنوال} + \frac{\text{ط ف}_١}{\text{ف}_١ + \text{ف}_٢}$$

$$= ١٦٥ + \frac{٥ \times ٦}{٦ + ٩} = ١٦٧ \text{ سنتيمتر}$$

خصائص المنوال:

١) يمتاز المنوال بالوضوح العملى لمعناه ولهذا يمكن استخدامه فى الحياة العملية فى حالات متعددة. فالمنوال هو القيمة الأكثر شيوعا، أى أن المنوال هو القيمة التى تتكرر وتتكرر فى الظاهرة بصورة أكثر من غيرها

ولهذا يمكن استخدام هذه القيمة المتكررة الشائعة فى الحياة العملية وعلـــى سبيل المثال فى الملبوسات بجميع أنواعها لجميع الأعمال للذكور والأناث

٢) لايخضع المنوال للعمليات الجبرية.

٣) يمتاز بالبساطة والسهولة فى طريقة حسابه كما يمكن ايجاده بالرسم.

٤) لايتأثر المنوال بالقيم الشاذة فى الظاهرة.

٥) قد لايوجد منوال لبعض الظواهر وذلك عندما لاتوجد قيمة متكررة .

٦) يجب استخدام التكرار المعدل عند ايجاد المنوال من التوزيعات التكرارية غير المتساوية الفئات وذلك حتى يمكن تحديد أكبر تكرار فى وحدة طـول الفئة.

وجدير بالذكر أن هناك عدة مقاييس أخرى للنزعة المركزيـــة مثـل الوسط الهندسى والوسط التوافقى والوسط التربيعى ولكنــها تعتــبر مقـاييس للموضع ذات طبيعة خاصة ولاتستخدم إلا فى حالة وجود أنواع معينـــة مـــن الظواهر والقيم وبالتالى فهى محدودة الاستخدام ويقتصر إستخدامها على هــذه القيم الخاصة، ولهذا سنكتفى بالمتوسطات الثلاثة السابق عرضها.

ثانيا : التشـــتت:

يعتبر إيجاد المتوسط لقيم الظاهرة الخطوة الأولى عند تحليـل الظـاهرة إحصائيا، وهذه الخطوة وإن كانت الخطوة الأساسـية الأولـى فـى التحليـل الاحصائى الا أنها بالتأكيد ليست كافية ويجب أن تتبعـها خطـوات متتاليـة للتحليل الاحصائى حتى تكتمل الصورة.

ويعتبر قياس التشتت فى الظاهرة الخطوة التالية فى التحليل الاحصـائى. ولتوضيح معنى التشتت فلنعتبر القيم التالية التى تعبر عن الوزن لأربع عينات عشوائية من الطلبة.

العينة الأولى: ٣٠، ٤٠، ٢٠، ٦٠، ٥٠

العينة الثانية: ٣٨، ٤٠، ٤١، ٤٢، ٣٩

العينة الثالثة: ٢٠، ٣٨، ٦٠، ٤٠، ٤٢

العينة الرابعة: ٤٠ ، ٤٠ ، ٤٠ ، ٤٠ ، ٤٠

وبإيجاد الوسط الحسابى والوسيط لكل عينة نجد أنه فى العينات الأربعـة تتساوى قيمة كل من الوسط الحسابى والوسيط لتساوى ٤٠ ولكن نلاحظ مـن الرسم التالى أن هناك اختلافا واضحا فى طبيعة قيم كل عينة حيث أن درجـة تقارب وتركز قيم مفردات العينة تختلف بين العينات الثلاثة حيث من الواضـح تركز قيـم مفردات العينة الثانية حول متوسطها بدرجة أكبر بكثير من تركـز قيم مفـردات العينة الأولى أو العينة الثالثة حول متوسطها بينما فـى العينـة الرابعة لايوجد أى اختلاف بين مفردات العينة وبين متوسطها، وعلـى هـذا يمكن القول أن:

	٦٠	٥٠	٤٠	٣٠	٢٠	١٠
العينة الأولى	•	•	•	•	•	
العينة الثانية			•••••			
العينة الثالثة	•		•••		•	
العينة الرابعة			•			

شكل رقم (٧)

وبتعبير آخر التشتت هو درجة التجانس بين قيم الظاهرة ودرجة التجمع والتركز لقيم الظاهرة حول متوسطها.

فاذا افترضنا أن لدينا القيم التالية:

٧٠، ٧٠، ٧٠، ٧٠، ٧٠، ٧٠

فاننا نقول أن هذه القيم متطابقة ويكون وسطها الحسابى يساوى ٧٠ ومن ثم فلايوجد أى تشتت لهذه القيم نظرا لعدم وجود أى فرق بين أى مفردة وبين الوسط الحسابى.

ولهذا كان من الضرورى فى التحليل الاحصائى قياس مدى تجانس قيم الظاهرة ومدى تركزها حول المتوسط وبتعبير آخر من الضرورى قياس التشتت للظاهرة كخطوة تالية فى التحليل الاحصائى.

فاذا اعتبرنا أن التشتت هو درجة تركز القيم حول متوسطها فإن طرق قياس التشتت تدور حول هذا المفهوم.

مقاييس التشتت

أولا : المدى:

يمكن تعريف المدى بأنه الفرق بين أكبر قيمة فى الظاهرة وأصغر قيمة فى الظاهرة، أى أن:

المدى = أكبر قيمة ــ أصغر قيمة ....(١٢)

ويتميز المدى بالبساطة والسهولة الشديدة عند حسابه، نلك لأن حساب المدى لايتطلب أكثر من إيجاد الفرق بين أكبر وأصغر قيم الظاهرة ويلعب المدى دورا كبيرا فى خرائط مراقبة الجودة فى العمليات الانتاجية حيث أنه من الطبيعى أن تكون هناك حدودا للجودة موضوعة مسبقا بحيث يكون الفرق بين الحد الأعلى والحد الأدنى هو المدى لجودة المنتج. أيضا نلاحظ أن كثيرا من التحاليل والاختبارات الطبية للانسان توجد فى شكل حدين أعلى وأدنى وحيث يمثل الفرق بينهما المدى الذى يتحرك فيه نتيجة التحليل للشخص السليم فإذا خرجت نتيجة التحليل عن أحد الحدود قد يدل نلك على وجود خلل طبى معين عند نلك الانسان.

ويعتبر المدى مقياسا سريعا وبسيطا لقياس تشتت الظاهرة، فإذا أردنا قياس المدى لبيانات الوزن للعينات العشوائية السابقة فإن :

المدى للعينة الأولى = ٦٠ - ٢٠ = ٤٠ كيلو جرام

المدى للعينة الثانية = ٤٢ - ٣٨ = ٤ كيلو جرام

المدى للعينة الثالثة = ٦٠ - ٢٠ = ٤٠ كيلو جرام

المدى للعينة الرابعة = ٤٠ - ٤٠ = صفر

وبنفس المفهوم يمكن إيجاد المدى من التوزيع التكرارى للظـــاهرة، فــإذا اعتبرنا التوزيع التكرارى التالى الذى يمثل توزيع مجموعة من الطلبة حســب فئات العمر.

فئات العمر	١٥-	١٨-	٢١-	٢٤-	٢٧-٣٠
عدد الطلبة	٦	٣٤	٥٠	٢٠	١٠

فيكون :

المدى = أكبر قيمة - أصغر قيمة

= ٣٠ - ١٥ = ١٥ كيلو جرام

عيوب المدى:

رغم بساطة وسهولة المدى فى الحساب الا أنه يعاب عليه مايلى:

١) يتأثر المدى وبشدة بالقيم الشاذة المتطرفة ولهذا يجــب تجنــب اســتخدام المدى كمقياس للتشتت اذا ماوجدت قيم شاذة ضمن قيم الظاهرة حيث أنه من المؤكد أن هذه القيمة الشاذة سوف تتسبب فى أن يكون المدى مضللا للتشتت ولنعتبر القيم ١٥، ٢٣، ١٨، ٢٤، ٢٠، ٨٥، ١٩.

فان المدى = أكبر قيمة ــ أصغر قيمة

= ٨٥ ــ ١٥ = ٧٠

وواضح أن القيمة ٨٥ هى قيمة متطرفة بالنسبة لباقى القيم، فــاذا كــانت هذه القيمة المتطرفة غير موجودة فان:

المدى = ٢٤ ــ ١٥ = ٩

٢) يأخذ المدى فى الحسبان قيمتين فقط هما أكبر قيمة وأصغر قيمة ويــهمل تماما باقى القيم أيا كان عددها أو قيمتها، فاذا اخترنا عينة عشوائية مــن

مائة طالب لدراسة ظاهرة الوزن وأردنا قياس التشتت من خلال المدى فإننا سنبحث فقط عن أكبر قيمة وأصغر قيمة وسنهمل تماما جميع القيم الأخرى وعددها ٩٨ قيمة.

ومن الطبيعي أن نقول أنه لاشك أن المقياس الاحصائي الذى يستخدم ويعالج رياضيا من خلال جميع قيم الظاهرة يعتبر أكثر كفاءة ودقة من المقياس الاحصائي الذى يهمل عددا كبيرا من قيم الظاهرة.

ثانيا: المدى الربيعي ونصف المدى الربيعي:

يمكننا التخلص من تأثر المدى بالقيم الشاذة المتطرفة باختيار المدى الربيعي كمقياس للتشتت باعتبار أن:

المدى الربيعي = الربيع الثالث ــ الربيع الأول

وباعتبار أن:

قيمة الربيع الثالث = $ب_٣$

قيمة الربيع الأول = $ب_١$

فان:

المدى الربيعي = $ب_٣$ ــ $ب_١$ ....(٨٣)

ونكون بذلك قد استبعدنا ربع القيم الأولى وربع القيم الأخيرة من قيم الظاهرة المرتبة ترتيبا تصاعديا أو ترتيبا تنازليا وبذلك يكون قد تم استبعاد القيم المتطرفة عند قياس التشتت.

فالربيع الأول والوسيط والربيع الثالث تقسم قيم الظاهرة الى أربعة أقسام، وقياس التشتت باستخدام المدى الربيعي يعنى أننا قد استبعدنا جزء من

المشاهدات من كل جانب ثم تم حساب المدى للقيم الباقية حيث أصغر قيمة هى الربيع الأول وأكبر قيمة متبقية هى الربيع الثالث. فكأن المدى الربيعى مقياسا يصف التشتت فى النصف الأوسط من المشاهدات ولهذا يرى كثير من الاحصائيين أخذ نصف المدى الربيعى باعتباره مقياسا للتشتت أفضل من المدى الربيعى. وبذلك يكون:

$$\text{نصف المدى الربيعى} = \frac{(\text{الربيع الثالث} - \text{الربيع الأول})}{٢}$$

$$\text{أى أن نصف المدى الربيعى} = \frac{(ب_٣ - ب_١)}{٢} \quad \text{.............................(١٤)}$$

وعادة مايتم إيجاد الوسيط كمقياسا للنزعة المركزية ومتوسط لقيم الظاهرة عند استخدام المدى الربيعى أو نصف المدى الربيعى كمقياس للتشتت.

مثال (٢١):

يوضح الجدول التكرارى التالى توزيع عينة عشوائية من ١٦٠ طالب على درجات الذكاء المختلفة. والمطلوب:

(أ) إيجاد كل من الوسيط والربيع الأول والربيع الثالث بالحساب وبالرسم.

(ب) قياس تشتت الذكاء.

الذكـــاء	٦٢-	٧٢-	٨٢-	٩٢-	١٠٢-	١١٢-١٢٢
عدد الطلبة	١٢	٢٨	٥٠	٤٠	٢٥	٥

الحـــل

كما أوضحنا سابقا فإنه يجب إعداد التوزيع التكرارى المتجمع الصاعد أو التوزيع التكرارى المتجمع النازل وذلك عند إيجاد الوسيط أو الربيع الأول أو الربيع الثالث. وبتكوين التوزيع الصاعد فيما يلى نجده كما يلى:

أقل من حدود الفئات	التكرار المتجمع الصاعد
أقل من ٦٢	صفر
أقل من ٧٢	١٢
أقل من ٨٢	٤٠
أقل من ٩٢	٩٠
أقل من ١٠٢	١٣٠
أقل من ١١٢	١٥٥
أقل من ١٢٢	١٦٠

$$\text{رتبة الوسيط} = \frac{\text{مجـ ك}}{٢} = \frac{١٦٠}{٢} = ٨٠$$

ونجد أن هذه الرتبة تقع بين ٤٠، ٩٠ ومن ثم يكون تكرار الفئة الوسيطية

= ٩٠ − ٤٠ = ٥٠

كما أن بداية فئة الوسيط = ٨٢ وطول فئة الوسيط = ط = ١٠

$ك_١ = ٤٠$

رتبة الوسيط = $ك_٢ = ٨٠$

تكرار الفئة الوسيطية = ك = ٥٠

$$\text{اذا قيمة الوسيط} = \text{بداية فئة الوسيط} + \frac{ط\,(ك_٢ - ك_١)}{ك_٣ - ك_١}$$

$$= ٨٢ + \frac{١٠\,(٨٠ - ٤٠)}{٩٠ - ٤٠}$$

= ٨٢ + ٨ = ٩٠ درجة

ويمكن إيجاد قيمة كل من الربيع الأول والربيع الثالث بنفس الطريقة التى أوجدنا بها قيمة الوسيط وذلك كما يلى:-

رتبة الربيع الأول = $\frac{مجـ ك}{٤} = \frac{١٦٠}{٤} = ٤٠$

وبالبحث عن رتبة الربيع الأول بين التكرارات المتجمعة الصاعدة نجد أنـــها موجودة مباشرة بين التكرارات المتجمعة مما يعنى أن قيمة الربيع الأول يمكن الحصول عليها مباشرة من التوزيع التصاعدى وهى تساوى ٨٢.

أى أن قيمة الربيع الأول = ٨٢ درجة

رتبة الربيع الثالث= $\frac{٣ مجـ ك}{٤} = \frac{٣ \times ١٦٠}{٤} = ١٢٠$

وبالبحث عن رتبة الربيع الثالث بين التكرارات المتجمعة نجد أنها تقع بين ٩٠ ، ١٣٠ مما يعنى أن تكرار فئة الربيع الثالث = ١٣٠ - ٩٠ = ٤٠

كما أن بداية فئة الربيع الثالث = ٩٢ وطول فئة الربيع الثالث = ١٠

اذا :

$ك_١ = ٩٠$

رتبة الربيع الثالث = $ك_٢ = ١٢٠$

تكرار فئة الربيع الثالث = ك = ٤٠

قيمة الربيع الثالث = بداية فئة الربيع الثالث + $\frac{ط (ك_٢ - ك_١)}{ك_٣ - ك_١}$

$= ٩٢ + \frac{١٠ (١٢٠ - ٩٠)}{١٣٠ - ٩٠}$

= ٩٢ + ٧,٥ = ٩٩,٥ درجة

فاذا استخدمنا المدى الربيعى كمقياس للتشتت فإن:

المدى الربيعى = الربيع الثالث – الربيع الأول

= $ب_٣ - ب_١$

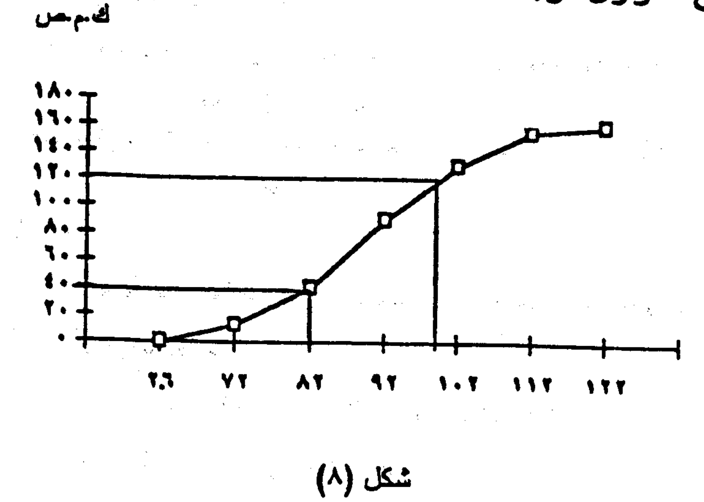
= ٩٩,٥ – ٨٢ = ١٧,٥ درجة

واذا استخدمنا نصف المدى الربيعى كمقياس للتشتت وهذا هو الأفضل فإن:

$$\text{نصف المدى الربيعى} = \frac{\text{الربيع الثالث} - \text{الربيع الأول}}{٢} = \frac{ب_٣ - ب_١}{٢}$$

$$= \frac{٩٩,٥ - ٨٢}{٢} = ٨,٧٥ \text{ درجة}$$

ويمكن إيجاد قيمة كل من الربيع الأول والربيع الثالث من الرسم وذلك برسم المنحنى المتجمع الصاعد أو المنحنى المتجمع النازل. ويوضح الشكل التالى إيجاد كلا من الربيع الأول والربيع الثالث من المنحنى المتجمع الصاعد للتوزيع التكرارى لدرجات الذكاء:



شكل (٨)

ويمكن القول أن استخدام الربيع الثالث والربيع الأول عند قياس التشتت لقيم الظاهرة يجنبنا وبدرجة كبيرة التأثر بالقيم المتطرفة ومن ثم فيعتبر المدى الربيعى ونصف المدى الربيعى كمقياس للتشتت أفضل من المدى وذلك بالاضافة الى امكانية إيجاد قيمة كل من الربيع الأول والربيع الثالث بالرسم ولكن ذلك لايمنع من القول أن استخدام المدى الربيعى أو نصف المدى الربيعى كمقياس للتشتت يعنى أننا نعتمد على قيمتين فقط (من بين جميع قيم الظاهرة) فى قياس التشتت ولانأخذ فى الاعتبار باقى قيم الظاهرة مما يعنى فقدان هذا المقياس لمبدأ الشمول الرياضى لقيم الظاهرة.

ثالثا: الانحراف المتوسط (عن الوسط الحسابى):

يمكن تعريف الانحراف المتوسط عن الوسط الحسابى بأنه متوسط الانحرافات المطلقة للقيم عن الوسط الحسابى.

فاذا كانت لدينا القيم غير المبوبة:

$س_1$ ، $س_2$ ، $س_3$ ، ..... $س_ن$ وكان الوسط الحسابى لهذه القيم هو $\bar{س}$ فإن:

$$\text{الانحراف المتوسط} = \frac{\text{مجـ} |س - \bar{س}|}{ن} \quad \text{.............} (١٥)$$

مثال (٢٢):

أوضحت عينة عشوائية من سبعة طلاب من طلبة الدراسات العليا بكلية التجارة أن أعمارهم كما يلى:

٢٥، ٢٠، ٢٨، ٣٢، ٤٠، ٢٤، ٢٧

أوجد الانحراف المتوسط عن الوسط الحسابى

الحـــل

بإيجاد الوسط الحسابى نجد أن:

$$\bar{\text{س}} = \frac{\text{مجـ س}}{\text{ن}} = \frac{١٩٦}{٧} = ٢٨ \text{ سنة}$$

الانحراف المتوسط عن الوسط الحسابى = مجـ | س — $\bar{\text{س}}$ | ÷ ن

= (|٢٥ — ٢٨| + |٢٠ — ٢٨| + |٢٨ — ٢٨| + |٣٢ — ٢٨| + |٤٠ — ٢٨| + |٢٤ — ٢٨| + |٢٧ — ٢٨|) ÷ ٧

$$= \frac{(٣ + ٨ + \text{صفر} + ٤ + ١٢ + ٤ + ١)}{٧} = \frac{٣٢}{٧} = ٤,٥٦ \text{ سنة}$$

ويمكن إيجاد الانحراف المتوسط من التوزيع التكرارى كما يلى :

١) نوجد الوسط الحسابى للتوزيع التكرارى.

٢) نوجد الانحراف المطلق (الفرق المطلق) بين الوسط الحسابى وبين مركز كل فئة من الفئات، أى | س — $\bar{\text{س}}$ |

٣) نقوم بترجيح الانحراف المطلق الذى حصلنا عليه لكل فئة بتكرار تلك الفئة أى نوجد | س — $\bar{\text{س}}$ | ك ويتم جمع حواصل الضرب السابقة أى مجـ | س — $\bar{\text{س}}$ | ك

٤) ويكون الانحراف المتوسط عن الوسط الحسابى هو حاصل قسمة مجموع حواصل الضرب السابقة على مجموع التكرارات، أى أن:

$$\text{الانحراف المتوسط} = \frac{\text{مجـ } | \text{س} - \bar{\text{س}} | \text{ ك}}{\text{مجـ ك}} \qquad \text{..........:.......(١٦)}$$

مثال (٢٣):

يوضح الجدول التالى التوزيع التكرارى للدخل الشهرى لعينة عشوائية من ١٠٠ طالب

الدخـــل	١٠-	١٦-	٢٢-	٢٨-	٣٤-٤٠
عدد الطلبة	٥	٢٥	٤٠	٢٠	١٠

والمطلوب إيجاد الانحراف المتوسط عن الوسط الحسابى

الحـــل

بإيجاد الوسط الحسابى للدخل الشهرى وبإتباع الخطوات السابقة يمكننا إعداد الجدول التالى:

الدخل	ك	س	س ك	|س - $\bar{س}$|	|س-$\bar{س}$| ك
١٠-	٥	١٣	٦٥	١٢,٣	٦١,٥
١٦-	٢٥	١٩	٤٧٥	٦,٣	١٥٧,٥
٢٢-	٤٠	٢٥	١٠٠٠	٠,٣	١٢,٠
٢٨-	٢٠	٣١	٦٢٠	٥,٧	١١٤,٠
٣٤-٤٠	١٠	٣٧	٣٧٠	١١,٧	١١٧,٠
المجموع	١٠٠		٢٥٣٠		٤٦٢,٠

$$\text{الوسط الحسابى} = \bar{س} = \frac{\text{مجـ س ك}}{\text{مجـ ك}} = \frac{٢٥٣٠}{١٠٠} = ٢٥,٣ \text{ جنيه}$$

ومن العلاقة (١٦) نجد أن:

$$\text{الانحراف المتوسط} = \frac{\text{مجـ }|س - \bar{س}|\text{ ك}}{\text{مجـ ك}} = \frac{٤٦٢}{١٠٠} = ٤,٦٢ \text{ جنيه}$$

ويتميز الانحراف المتوسط بصفة عامة بأنه يأخذ فى الحسبان مختلف القيم إلا أنه يعاب عليه أنه يتأثر بالقيم المتطرفة ولايمكن حسابه من الجداول التكرارية المفتوحة، أيضا يعاب عليه رياضيا إستخدام الانحرافات المطلقة، فجميع الانحرافات موجبة أيا كانت حقيقتها.

رابعا: الانحراف المعيارى:

بفرض أن لدينا عينة عشوائية من ن مفردة من القيم $س_١، س_٢، .. س_ن$ وحيث متوسطها $\bar{س}$ فإن انحرافات القيم عن المتوسط $\bar{س}$ يكون:
$(س_١ - \bar{س})، (س_٢ - \bar{س})، ....، (س_ن - \bar{س})$ وحيث بعض هذه الانحرافات موجب والبعض الأخر سالب ومجموع الانحرافات يساوى الصفر. ولقياس التشتت بإستخدام الانحراف المعيارى فإننا نتخلص من الاشارات السالبة باسلوب رياضى سليم وهو إيجاد مربع تلك الانحرافات وبذلك يمكن ايجاد التشتت باستخدام الانحراف المعيارى :
حيث أن الانحراف المعيارى هو الجذر التربيعى لمتوسط مربع انحرافات القيم عن وسطها الحسابى

أى أن الانحراف المعيارى = $$\sqrt{\frac{مجـ (س - \bar{س})^٢}{ن}} \quad ....(١٧)$$

ويلاحظ أن وحدات القياس للانحراف المعيارى هى نفس وحدة القياس للظاهرة.

وفى الحقيقة فإنه يجب القسمة على (ن - ١) حيث تعطينا النظرية الاحصائية الأساس والتفسير العلمى لاستخدام (ن - ١) بدلا من ن حيث أنه عند استخدام (ن - ١) فى التقدير فإن الانحراف المعيارى للعينة يعتبر تقديرا غير متحيزا للانحراف المعيارى للمجتمع ولكننا للبساطة سنستخدم دائما ن بدلا من ن - ١

مثال (٢٤):

مجتمع يتكون من خمسة مفردات هى:

٢٢، ١٨، ٢٤، ٢٠، ٢٦ سنة

أوجد الانحراف المعيارى.

الحـــل

$$\text{متوسط المجتمع} = \text{م} = \frac{(١٨ + ٢٢ + ٢٤ + ٢٠ + ٢٦)}{٥}$$

م = ٢٢ سنة

$$\text{الانحراف المعيارى} = \sqrt{\frac{\text{مجـ}(\text{س} - \text{م})^{٢}}{\text{ن}}}$$

$$= \sqrt{\frac{\left[(١٨-٢٢)^{٢} + (٢٢-٢٢)^{٢} + (٢٤-٢٢)^{٢} + (٢٠-٢٢)^{٢} + (٢٦-٢٢)^{٢}\right]}{٥}}$$

$$= \sqrt{\frac{(١٦ + \text{صفر} + ٤ + ٤ + ١٦)}{٥}} = \sqrt{٨} = ٢{,}٨٣ \text{ سنة}$$

ويعتبر الانحراف المعيارى أكثر مقاييس التشتت استخداما فى النظرية الاحصائية وفى الاستدلال الاحصائى وذلك لما يتميز به من خصائص احصائية تؤهله لذلك ولكن قد يعاب عليه تأثره بالقيم المتطرفة.

مثال (٢٥):

١ إختيرت عينة عشوائية من خمسة طلاب فكانت أعمـــارهم ٢٢، ٢٣، ٢١، ٢٤، ٢٥ سنة أوجد الانحراف المعيارى للعمر.

الحـــل

$$\overline{س} = \frac{مجـ س}{ن} = \frac{١١٥}{٥} = ٢٣ \text{ سنة}$$

$$\text{الانحراف المعيارى} = \sqrt{\frac{مجـ \left(س - \overline{س}\right)^{٢}}{ن}} = \sqrt{\frac{١+٠+٤+١+٤}{٥}}$$

$$= \sqrt{٢} = ١,٤ \text{ سنة}$$

الانحراف المعيارى من بيانات مبوبة:

يمكن إيجاد الانحراف المعيارى للظاهرة من البيانات المبوبة فى جـــدول تكرارى بتطبيق نفس المفهوم السابق عرضه وذلك بإعتبار أن التبـــاين هـــو متوسط مربعات انحرافات القيم عن الوسط الحسابى مع ملاحظة أن:

١) نعتبر أن مركز الفئة هو القيمة التى تمثل قيمة الفئة ومن ثم يجب إيجـــاد الانحراف والفرق بين الوسط الحسابى وبين مركز كل فئة.

٢) يتم ترجيح مربع الانحراف السابق فى كل فئة بتكرار الفئة.

ويكون الانحراف المعيارى هو متوسط مربعات الانحرافات المرجحة الســـابقة بعد إيجاد الجذر التربيعى لها

أى أن.

$$\text{الانحراف المعيارى} = \sqrt{\frac{مجـ ك \left(س - \overline{س}\right)^{٢}}{(مجـ ك - ١)}} \quad \text{.......(١٨)}$$

ولتبسيط العمل الحسابى فإنه سيلاحظ دائما أنه عند إيجاد التباين أو الانحراف المعيارى سنقوم بالقسمة على مجـ ك فقط أى أننا سنعتبر أن :

$$\text{الانحراف المعيارى} = \sqrt{\frac{\text{مجـ ك (س} - \bar{\text{س}})^2}{\text{مجـ ك}}} \quad \ldots\ldots\ldots\ldots(١٩)$$

مثال (٢٦):

أوجد الانحراف المعيارى للتوزيع التكرارى التالى الذى يمثل توزيع عينة من ٨٠ طالب على فئات العمر المختلفة.

العمر	١٥-	١٧-	١٩-	٢١-	٢٣-٢٥
عدد الطلبة	٦	١٤	٤٢	١٠	٨

الحل

حيث أننا نحصل على الانحراف المعيارى باعتباره الجذر التربيعى لمتوسط مربعات انحرافات القيم عن الوسط الحسابى فإننا نبدأ بإيجاد الوسط الحسابى:

العمر	ك	س	س ك	(س − $\bar{\text{س}}$)	ك(س − $\bar{\text{س}}$)^2
١٥-	٦	١٦	٩٦	− ٤	٩٦
١٧-	١٤	١٨	٢٥٢	− ٢	٥٦
١٩-	٤٢	٢٠	٨٤٠	صفر	صفر
٢١-	١٠	٢٢	٢٢٠	+ ٢	٤٠
٢٣-٢٥	٨	٢٤	١٩٢	+ ٤	١٢٨
المجموع	٨٠		١٦٠٠		٣٢٠

الوسط الحسابى = $\overline{س}$ = $\frac{مجـ س ك}{مجـ ك}$ = $\frac{١٦٠٠}{٨٠}$ = ٢٠ سنة

ويكون الانحراف المعيارى = $\sqrt{\frac{مجـ ك (س - \overline{س})^{٢}}{مجـ ك}}$ = $\sqrt{\frac{٣٢٠}{٨٠}}$ = ٢ سنة

ولكن قد نجد صعوبة فى إيجاد مربع انحراف القيم عن الوسط الحسابى وذلك اذا كانت قيمة الوسط الحسابى بها كسور، ولهذا يمكن تبسيط العمل الحسابى على اعتبار أن:

$$مجـ (س - \overline{س})^{٢} = مجـ س^{٢} - ن (\overline{س})^{٢}$$

كما أن: $مجـ ك (س - \overline{س})^{٢} = مجـ س^{٢} ك - مجـ ك (\overline{س})^{٢}$

وبذلك يمكن إستخدام الصيغ التالية عند إيجاد الانحراف المعيارى مع ملاحظة أننا سنستخدم دائما للبساطة ن بدلا من (ن – ١) للبيانات الغير مبوبة وذلك طالما أننا لن نستخدم التباين أو الانحراف المعيارى فى الاستدلال الاحصائى فى هذه المرحلة.

الانحراف المعيارى = $\sqrt{\frac{مجـ س^{٢}}{ن} - \overline{س}^{٢}}$ ....(٢٠)

كما أن الانحراف المعيارى للبيانات المبوبة سيكون فى الصورة

الانحراف المعيارى = $\sqrt{\frac{مجـ س^{٢} ك}{مجـ ك} - \overline{س}^{٢}}$ ....(٢١)

مثال (٢٧):

أوجد الانحراف المعيارى للقيم التالية:

٧ ، ١٠ ، ٤ ، ٨ ، ١٢ ، ١٣

الحــــل

الوسط الحسابى = $\bar{س}$ = $\frac{مجـ س}{ن}$ = $\frac{٧ + ١٠ + ٤ + ٨ + ١٢ + ١٣}{٦}$

= $\frac{٥٤}{٦}$ = ٩

كما يجب إيجاد مجـ $س^2$ لإيجاد التباين:

مجـ $س^2$ = ٤٩ + ١٠٠ + ١٦ + ٦٤ + ١٤٤ + ١٦٩

= ٥٤٢

ومن المعادلة (٢٠)

∴ الانحراف المعيارى = $\sqrt{\frac{مجـ س^2}{ن} - \bar{س}^2}$

= $\sqrt{\frac{٥٤٢}{٦} - (٩)^2}$

= $\sqrt{٩٠,٣ - ٨١}$ = $\sqrt{٩,٣}$ = ٣,١

ولعل من الأفضل تنظيم العمل الحسابى كما يلى:

س٢	س
٤٩	٧
١٠٠	١٠
١٦	٤
٦٤	٨
١٤٤	١٢
١٦٩	١٣
٥٤٢	٥٤

مثال (٢٨):

أوجد الانحراف المعيارى للتوزيع التكرارى فى مثـــال (٢٦) باســتخدام مراكز الفئات مباشرة

الحـــل

بإيجاد الوسط الحسابى وتطبيق المعادلة (٢١) نجد أن:

العمــر	ك	س	س ك	س٢ ك
١٥-	٦	١٦	٩٦	١٥٣٦
١٧-	١٤	١٨	٢٥٢	٤٥٣٦
١٩-	٤٢	٢٠	٨٤٠	١٦٨٠٠
٢١-	١٠	٢٢	٢٢٠	٤٨٤٠
٢٣-٢٥	٨	٢٤	١٩٢	٤٦٠٨
المجموع	٨٠		١٦٠٠	٣٢٣٢٠

$$\bar{س} = \frac{مجـ\ س\ ك}{مجـ\ ك} = \frac{١٦٠٠}{٨٠} = ٢٠\ سنة$$

$$\text{الانحراف المعيارى} = \sqrt{\frac{\text{مجـ س}^2\text{ ك}}{\text{مجـ ك}} - \bar{\text{س}}^2}$$

$$= \sqrt{\frac{٣٢٣٢٠}{٨٠} - ٤٠٠} = ٢ \text{ (سنة)}$$

وكما أوضحنا عند مناقشة الوسط الحسابى فإنه يمكن تبسيط مراكز الفئات بطرح وسط فرضى من جميع مراكز الفئات وبذلك نحصل على انحرافات مراكز الفئات عن الوسط الفرضى (ح) وتستخدم هذه الانحرافات بعد ترجيحها بالتكرارات فى إيجاد كلا من الوسط الحسابى والانحراف المعيارى بحيث يصبح الانحراف المعيارى فى الصورة:

$$\text{الانحراف المعيارى} = \text{ع} = \sqrt{\frac{\text{مجـ ح}^2\text{ ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right)^2} \quad ....(٢٢)$$

مثال (٢٩):

أوجد الانحراف المعيارى من التوزيع التكرارى التالى بإستخدام طريقة الانحراف عن وسط فرضى.

الفئـــات	٢٥-	٢٨-	٣١-	٣٤-	٣٧-٤٠
ك	٤	١٦	٣٥	٢٥	٥

الحـــل

باعداد الجدول التالى حيث الوسط الفرضى = ٣٢,٥

الفئات	ك	س	ح	ح ك	ح^{٢} ك
-٢٥	٤	٢٦,٥	- ٦	- ٢٤	١٤٤
-٢٨	١٦	٢٩,٥	- ٣	- ٤٨	١٤٤
-٣١	٣٥	٣٢,٥	صفر	صفر	صفر
-٣٤	٢٥	٣٥,٥	+ ٣	+ ٧٥	٢٢٥
٤٠-٣٧	٥	٣٨,٥	+ ٦	+ ٣٠	١٨٠
المجموع	٨٥			+ ٣٣	٦٩٣

لانحراف المعيارى = ع = $\sqrt{\frac{\text{مجـ ح}^{٢}\text{ ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right)^{٢}}$

$= \sqrt{\frac{٦٩٣}{٨٥} - \left(\frac{٣٣}{٨٥}\right)^{٢}}$

$= \sqrt{٨,١٥٢٩ - ٠,١٥٠٧}$

$= ٢,٨٣$

ويمكن أيضا تبسيط الانحرافات بالقسمة على مقدار ثابت (غالبا مـا يكون هذا المقدار الثابت هو نفسه طول الفئة للجدول المنتظـم المتسـاوى الفئـات وسنرمز للمقدار الثابت بالرمز طـ ويصبح حينئذ الانحراف المعيارى فـى الصورة:

الانحراف المعيارى = ع = طـ × $\sqrt{\frac{\text{مجـ ح}^{٢}\text{ ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right)^{٢}}$ ....(٢٣)

مثال (٣٠):

أوجد الانحراف المعيارى من التوزيع التكرارى التالى باستخدام طريقة الانحراف المبسطة.

الدخل	-٢٠٠	-٢٥٠	-٣٠٠	-٣٥٠	٤٠٠-٤٥٠
ك	٦	١٤	٤٠	٢٥	١٥

الحل

باعداد الجدول التالى

الدخل	ك	س	ح	ح'	ح' ك	ح'٢ ك
-٢٠٠	٦	٢٢٥	-١٠٠	-٢	-١٢	٢٤
-٢٥٠	١٤	٢٧٥	-٥٠	-١	-١٤	١٤
-٣٠٠	٤٠	٣٢٥	صفر	صفر	صفر	صفر
-٣٥٠	٢٥	٣٧٥	+٥٠	+١	+٢٥	٢٥
-٤٠٠	١٥	٤٢٥	+١٠٠	+٢	+٣٠	٦٠
المجموع	١٠٠				+٢٩	١٢٣

ومن العلاقة (٢٣)

$$\text{الانحراف المعيارى} = ع = ط \times \sqrt{\frac{\text{مج ح}'^{2}\text{ك}}{\text{مج ك}} - \left(\frac{\text{مج ح}'\text{ك}}{\text{مج ك}}\right)^{٢}}$$

$$ع = ٥٠ \times \sqrt{\frac{١٢٣}{١٠٠} - \left(\frac{٢٩}{١٠٠}\right)^{٢}}$$

$$= ٥٠ \times \sqrt{١,٢٣ - ٠,٠٨٤١}$$

$$= ٥٣,٥$$

ومن البديهى أنه يمكن إيجاد الانحراف المعيارى بالاستفادة من الطريقة التى نوجد بها أولا: الوسط الحسابى فإذا استخدمنا مراكز الفئات مباشرة فى إيجاد الوسط الحسابى فإننا أيضا نستخدم مراكز الفئات مباشرة لإيجاد الانحراف المعيارى ونستخدم العلاقة (٢١) لإيجاد قيمة الانحراف المعيارى.

وإذا اخترنا وسطا فرضيا من مراكز الفئات وتم طرح الوسط الفرضى من جميع مراكز الفئات للحصول على انحرافات مراكز الفئات عن الوسط الفرضى (ح) فإننا أيضا نستخدم (ح) فى إيجاد الانحراف المعيارى ونستخدم العلاقة (٢٢) فى إيجاد الانحراف المعيارى.

وإذا اختزلنا الانحرافات (ح) بقسمتها على مقدار ثابت (ط) وحصلنا بذلك على الانحرافات المختزلة (وذلك لإيجاد الوسط الحسابى) فإننا أيضا نستخدم الانحرافات المختزلة فى إيجاد الانحراف المعيارى بإستخدام العلاقة (٢٣).

وجدير بالذكر أن الانحراف المعيارى أكثر مقاييس التشتت إستخداما كما أنه أكثرهم اتساقا وتفسيرا ومنطقا للتشتت فنحن نقيس انحراف وتشتت كل مفردة فى قيم الظاهرة عن الوسط الحسابى لتلك الظاهرة، ونظرا لأن من خواص الوسط الحسابى أن مجموع انحرافات القيم عنه يساوى الصفر فقد كان من الضرورى التخلص من الاشارات السالبة فى انحرافات القيم عن وسطها الحسابى وكان الاسلوب الرياضى الأفضل هو تربيع جميع الانحرافات لتكون كلها ذات اشارات موجبة، وبأخذ متوسط تلك الانحرافات المربعة نحصل على متوسط مربعات انحرافات القيم عن الوسط الحسابى (التباين) ولكن وحدة القياس للتباين تكون مربعة كما أن الانحرافات كلها انحرافات مربعة لهذا يتم أخذ الجذر التربيعى للتباين لتكون وحدة القياس للانحراف

المعيارى هى نفس وحدة القياس للظاهرة. كما يمتاز الانحراف المعيارى باعتماده على جميع قيم الظاهرة ولهذا فهو أفضل من المدى الربيعى والمدى فى هذا الشأن .

ويتميز الانحراف المعيارى أيضا بأن قيمته لاتتأثر مطلقا اذا طرحنا مقدارا ثابتا من جميع قيم الظاهرة لتبسيط العمليات الحسابية. ولتوضيح تلك الخاصية فلنعتبر القيم التالية:

س : ٩٠٤، ٩٠٨، ٩١٠، ٩٠٢

$$\text{فإن الوسط الحسابى} = \frac{\text{مجـ س}}{\text{ن}}$$

= ٣٦٢٤ ÷ ٤ = ٩٠٦ وتكون انحرافات القيم عن الوسط الحسابى هى:

$(\text{س} - \overline{\text{س}}) = -٢ ، +٢ ، +٤ ، -٤$

$\text{اذا مجـ}(\text{س} - \overline{\text{س}})^{٢} = ٤ + ٤ + ١٦ + ١٦ = ٤٠$

$$\text{إذا الانحراف المعيارى} = \sqrt{\frac{\text{مجـ}\left(\text{س} - \overline{\text{س}}\right)^{٢}}{\text{ن}}}$$

$$= \sqrt{\frac{٤٠}{٤}} = \sqrt{١٠}$$

فاذا طرحنا من القيم السابقة أى رقم ثابت لتبسيط العمل الحسابى وليكن رقم ٩٠٠ على سبيل المثال فإن القيم الجديدة هى : ٤، ٨، ١٠، ٢

ويكون الوسط الحسابى الجديد = ٢٤ ÷ ٤ = ٦

$\text{ولكن سنجد أن الانحراف المعيارى} = \sqrt{(٤ + ٤ + ١٦ + ١٦) \div ٤}$

$= \sqrt{١٠}$

وهى نفس قيمة الانحراف المعيارى للقيم الأصلية، ولكن يعاب على الانحراف المعيارى تأثره بالقيم المتطرفة. ولهذا السبب يجب عدم استخدام التباين أو الانحراف المعيارى فى قياس التشتت اذا كانت هناك قيما متطرفة فى الظاهرة.

أيضا يعاب على الانحراف المعيارى عدم امكانية حسابه من الجداول التكرارية المفتوحة.

ويمكننا القول الأن بعد أن عرضنا لمقاييس التشتت المختلفة أنه عادة مايتم اختيار مقياس التشتت المناسب بناء على الاختيار المسبق لمتوسط الظاهرة، ونوعية التوزيع التكرارى للظاهرة.

فمعرفة الوسيط كمتوسط للظاهرة يجعل من المناسب استخدام نصف المدى الربيعى أو الانحراف المتوسط عن الوسيط للتعبير عن مقياس التشتت فى الظاهرة. واذا كان التوزيع التكرارى مفتوحا فى هذه الحالة فليس أمامنا من سبيل لقياس التشتت الا من خلال نصف المدى الربيعى.

وعند إيجاد الوسط الحسابى للتعبير عن متوسط الظاهرة فإن الانحراف المعيارى كمقياس للتشتت يساعدنا على الحكم على كفاءة الوسط الحسابى كمتوسط للظاهرة، وأيضا اذا ماكنا نرغب فى إستخدام بعض أساليب الاستدلال الاحصائى لمعلمات المجتمع فيجب أن نتعامل مع التباين والانحراف المعيارى.

واذا أردنا إعطاء أوزان متساوية لجميع انحرافات قيم الظاهرة عن متوسطها فإنه يمكننا حينئذ استخدام الانحراف المتوسط. ويمكننا استخدام المدى فى إيجاد مقياس سريع وبسيط للتشتت ولكن بشرط ألا يكون هناك اختلافا واسعا بين مختلف قيم الظاهرة.

ثالثا : مقاييس التشتت النسبى:

أوضحنا سابقا أن لمقياس التشتت وحدة قياس من نفس طبيعة وحدة قياس الظاهرة، ولذلك لانستطيع مقارنة التشتت بين ظاهرتين مختلفتين فى وحدة القياس. فليس من المتصور على سبيل المثال أن نقارن الانحراف المعيارى لأعمار طلبة كلية التجارة (وهو مقيس بالسنوات) بالانحراف المعيارى للأوزان لطلبة نفس الكلية (وهو مقيس بالكيلو جرام) حيث أنه لايمكن مقارنة سنوات مع كيلو جرامات على سبيل المثال، وهكذا لايمكننا مقارنة التشتت بين أى ظاهرتين مختلفتين فى وحدة القياس.

ويفرض علينا المنطق العلمى أيضا عدم المقارنة بين ظاهرتين لهما نفس وحدة القياس ولكن تختلفان فيما بينها فى متوسط كل منهما. وعلى سبيل المثال لايجب المقارنة بين الانحراف المعيارى للوزن للطلبة بكلية التجارة وبين الانحراف المعيارى للوزن للطالبات بكلية التجارة وذلك رغم أن كل منهما مقيس بالكيلو جرام ولكن لاتصح المقارنة فى هذه الحالة لأننا نقارن بين مجتمعين يختلفان فى المتوسط لكل منهما وفى طبيعة توزيع مفردات كل مجتمع حول هذا المتوسط.

لهذا كان من الضرورى عند مقارنة التشتت بين عدة ظواهر مراعاة:

أ) اختلاف وحدة القياس بين الظواهر.

ب) اختلاف المتوسط لكل ظاهرة وطبيعة توزيع المفردات حول المتوسط فى كل ظاهرة.

أو بتعبير آخر يمكننا مقارنة التشتت بين الظواهر المختلفة من خـــلال إيجاد التشتت النسبى للظاهرة والذى يعبر عن التشتت فى صورة نسبة مئويـــة وبدون وحدة القياس الخاصة بالظاهرة وذلك بإيجاد ماتطلق عليـــه احصائيـــا معامل الاختلاف.

ويعتمد تركيب معامل الاختلاف على كل من مقياس التشتت بالاضافة الى المتوسط الخاص به.

١ - فإذا استخدمنا المدى كمقياس للتشتت فإن معامل الاختلاف يكون فـــى الصورة :

$$\text{معامل الاختلاف} = \frac{\text{أكبر قيمة} - \text{أصغر قيمة}}{\text{أكبر قيمة} + \text{صغر قيمة}} \times ١٠٠ \quad ....(٢٤)$$

٢ - وإذا استخدمنا نصف المدى الربيعى كمقياس للتشـــتت فـــإن معـــامل الاختلاف يكون فى الصورة.

$$\text{معامل الاختلاف} = \frac{\text{نصف المدى الربيعى}}{\text{الوسيط}} \times ١٠٠ \quad ....(٢٥)$$

مثال (٣١):

يرغب أحد الباحثين فى مقارنة تشتت ظاهرتى الوزن والطول لطلبة كلية التجارة. وقد أوضحت عينة عشوائية من ١٠٠ طالب أن:

الوسيط للوزن = ٧٥ كيلو جرام

نصف المدى الربيعى للوزن = ١٢ كيلو جرام

بينما كان التوزيع التكرارى للطول كما يلى:

الطـــول	-١٥٠	-١٦٠	-١٧٠	-١٨٠	١٩٠-٢٠٠
عدد الطلبة	٥	٢٥	٤٥	٢٢	٣

قارن بين تشتت الطول وتشتت الوزن

الحـــل

كما أوضحنا فإننا لانستطيع المقارنة بين تشتت ظاهرتين الا من خـــلال مقياس نسبى للتشتت حتى يمكن التخلص من وحدة القياس المطلقة لكل ظاهرة وبإيجاد معامل الاختلاف لكل ظاهرة نجد أن:

$$\text{معامل الاختلاف} = \frac{\text{نصف المدى الربيعى}}{\text{الوسيط}} \times ١٠٠$$

$$\text{اذا معامل الاختلاف للوزن} = \frac{١٢}{٧٥} \times ١٠٠ = ١٦\%$$

ولإيجاد معامل الاختلاف للطول يجب إيجاد الوسيط والربيع الأول والربيـــع الثالث للتوزيع التكرارى للطول.

أقل من حدود الفئات	التكرار المتجمع الصاعد
أقل من ١٥٠	صفر
أقل من ١٦٠	٥
أقل من ١٧٠	٣٠
أقل من ١٨٠	٧٥
أقل من ١٩٠	٩٧
أقل من ٢٠٠	١٠٠

رتبة الوسيط = ١٠٠ + ٢ = ٥٠

وبالبحث عن رتبة الوسيط بين التكرارات المتجمعة نجد أنها محصورة بين ٣٠ ، ٧٥ ومن ثم يكون:

ك١ = ٣٠ ، ك٢ = ٥٠ ، ك٣ = ٧٥

وتكرار الفئة الوسيطية = ك = ٧٥ − ٣٠ = ٤٥

وقيمة الوسيط تتراوح بين ١٧٠، ١٨٠ ومن ثم طول فئة الوسيط = ١٠

الوسيط = بداية فئة الوسيط + ط (ك٢ − ك١) / (ك٣ − ك١)

= ١٧٠ + ١٠ (٥٠ − ٣٠) / (٧٥ − ٣٠)

= ١٧٠ + ٤.٤ = ١٧٤.٤ سم

رتبة الربيع الأول = مجـ ك / ٤ = ١٠٠ / ٤ = ٢٥

وبالبحث فى التوزيع التكرارى المتجمع الصاعد نجد أن رتبة الربيع الأول محصورة بين ٥، ٣٠ ومن ثم نجد أن:

ك١ = ٥ ، ك٢ = ٢٥ ، ك٣ = ٣٠

كما أن تكرار فئة الربيع الأول = ك = ٣٠ − ٥ = ٢٥

كما أن الربيع الأول يتراوح بين ١٦٠ ، ١٧٠

قيمة الربيع الأول = بداية فئة الربيع الأول + ط (ك٢ − ك١) / (ك٣ − ك١)

$$= 160 + \frac{10\,(25 - 5)}{30 - 5}$$

$= 160 + 8 = 168$ سم

رتبة الربيع الثالث = ٣ مجـ ك ÷ ٤

$$= \frac{3 \times 100}{4} = 75$$

وبالبحث عن رتبة الربيع الثالث نجد أنها توجـــد مباشـــرة فـــى التكـــرارات المتجمعة ومن ثم نجد أن:

قيمة الربيع الثالث تساوى ١٨٠ سم

$$\text{ويكون نصف المدى الربيعى} = \frac{(\text{الربيع الثالث} - \text{الربيع الأول})}{2}$$

$$= \frac{180 - 168}{2} = \frac{12}{2} = 6 \text{ سم}$$

$$\text{اذا معامل الاختلاف للطول} = \frac{\text{نصف المدى الربيعى}}{\text{الوسيط}} \times 100$$

$$= \frac{6}{174.4} \times 100$$

$= 3.4\ \%$

وبمقارنة معامل الاختلاف للوزن بمعامل الاختلاف للطول نستنتج أن توزيـــع الوزن أكثر تشتت من توزيع الطول.

٣ – واذا استخدمنا الانحراف المتوسط عن الوسط الحسابى كمقايس للتشـــتت فإن معامل الاختلاف يكون فى الصورة:

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المتوسط}}{\text{الوسط الحسابى}} \times ١٠٠ \quad ....(٢٦)$$

مثال (٣٢):

يوضح الجدول التالى التوزيع التكرارى لعينة عشوائية من ١٠٠ طالب موزعة حسب فئات الدخل الشهري بالجنيه.

الدخل الشهرى	٥-	١١-	١٧-	٢٣-	٢٩-٣٥
عدد الطلبة	٥	٢٥	٤٠	٢٠	١٠

بينما بدراسة العمر لنفس العينة كان الوسط الحسابى للعمر ٢٠ سنة والانحراف المتوسط عن الوسط الحسابى ٣ سنوات. والمطلوب:

أ - مقارنة تشتت العمر مع تشتت الدخل.

ب - إيجاد المدي ومعامل الاختلاف للدخل.

الحل

بإيجاد معامل الاختلاف لكل ظاهرة نجد أن :

$$\text{معامل الاختلاف للعمر} = \frac{\text{الانحراف المتوسط}}{\text{الوسط الحسابى}} \times ١٠٠$$

$$= \frac{٣}{٢٠} \times ١٠٠ = ١٥ \%$$

ولإيجاد معامل الاختلاف للدخل يجب إيجاد الوسط الحسابى للدخل والانحراف المتوسط عن الوسط الحسابى ولهذا يلزم اعداد الجدول التالى:

الدخل	ك	س	س ك	|س - س̄|	|س - س̄| ك
-٥	٥	٨	٤٠	١٢,٣	٦١,٥
-١١	٢٥	١٤	٣٥٠	٦,٣	١٥٧,٥
-١٧	٤٠	٢٠	٨٠٠	٠,٣	١٢,٠
-٢٣	٢٠	٢٦	٥٢٠	٥,٧	١١٤,٠
٢٩-٣٥	١٠	٣٢	٣٢٠	١١,٧	١١٧,٠
المجموع	١٠٠		٢٠٣٠		٤٦٢,٠

الوسط الحسابي للدخل = س̄ = $\frac{\text{مجـ س ك}}{\text{مجـ ك}}$

= ٢٠٣٠ ÷ ١٠٠ = ٢٠,٣٠ جنيه

الانحراف المتوسط عن الوسط الحسابي = $\frac{\text{مجـ |س - س̄| ك}}{\text{مجـ ك}}$

= ٤٦٢ ÷ ١٠٠ = ٤,٦٢ جنيه

معامل الاختلاف للدخل = $\frac{\text{الانحراف المتوسط}}{\text{س̄}}$ × ١٠٠

= $\frac{٤,٦٢}{٢٠,٣}$ × ١٠٠ = ٢٢,٨ %

وبمقارنة معامل الاختلاف لكل من العمر والدخل يتضح أن ظـــاهرة الدخـــل أكثر تشتتا من العمر.

من العلاقة (١٢)

المدى = أكبر قيمة – أصغر قيمة

$$٣ \text{ معامل الاختلاف} = \frac{\text{أكبر قيمة} - \text{أصغر قيمة}}{\text{أكبر قيمة} + \text{صغر قيمة}} \times ١٠٠$$

$$= \frac{٣٥ - ٥}{٣٥ + ٥} \times ١٠٠ = ٧٥\%$$

٤ - أيضا يمكننا إيجاد معامل الاختلاف بإستخدام الانحراف المعيارى والوسط الحسابى حيث:

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعيارى}}{\bar{س}} \times ١٠٠ \quad ....(٢٧)$$

$$٣٣ = \frac{ع}{\bar{س}} \times ١٠٠$$

مثال (٣٣):

أوجد الانحراف المعيارى للعمر من التوزيع التكرارى التالى واستخدمه فى إيجاد معامل الاختلاف.

فئات العمر	١٠-	١٨-	٢٦-	٣٤-	٤٢-٥٠
ك	٤	١٦	٣٠	٢٠	١٠

الحـــل

يمكن إعداد الجدول التالى لإيجاد الوسط الحسابى والانحراف المعيارى للعمر

العمـــر	ك	س	س ك	$س^٢$ ك
١٠-	٤	١٤	٥٦	٧٨٤
١٨-	١٦	٢٢	٣٥٢	٧٧٤٤
٢٦-	٣٠	٣٠	٩٠٠	٢٧٠٠٠
٣٤-	٢٠	٣٨	٧٦٠	٢٨٨٨٠
٤٢-٥٠	١٠	٤٦	٤٦٠	٢١١٦٠
المجموع	٨٠		٢٥٢٨	٨٥٥٦٨

$$\bar{\text{س}} = \frac{\text{مجـ س ك}}{\text{مجـ ك}}$$

$$= \text{٢٥٢٨} \div \text{٨٠} = \text{٣١,٦ سنة}$$

$$\text{الانحراف المعيارى} = \sqrt{\frac{\text{مجـ س}^{\text{٢}}\text{ ك}}{\text{مجـ س}} - (\bar{\text{س}})^{\text{٢}}}$$

$$= \sqrt{\frac{\text{٨٥٥٦٨}}{\text{٨٠}} - (\text{٣١,٦})^{\text{٢}}}$$

$$= \sqrt{\text{١٠٦٩,٦} - \text{٩٧٩,٦٩}} = \sqrt{\text{٨٩,٦١}}$$

$$= \text{٩,٥ سنة}$$

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعيارى}}{\bar{\text{س}}} \times \text{١٠٠}$$

$$= \frac{\text{٩,٥}}{\text{٣١,٦}} \times \text{١٠٠} = \text{٣٠\%}$$

مثال (٣٤):

أوجد الانحراف المتوسط عن الوسط الحسابى من التوزيع التـالى ثـم أوجد معامل الاختلاف.

العمر	١٠-	١٢-	١٤-	١٦-	١٨-٢٠	المجموع
ك	٣	٧	٢١	٥	٤	٤٠

الحـــل

العمر	ك	س	س ك	|س - $\bar{س}$|	|س - $\bar{س}$| ك
١٠-	٣	١١	٣٣	٤+	١٢
١٢-	٧	١٣	٩١	٢+	١٤
١٤-	٢١	١٥	٣١٥	صفر	صفر
١٦-	٥	١٧	٨٥	٢+	١٠
١٨-٢٠	٤	١٩	٧٦	٤+	١٦
المجموع	٤٠		٦٠٠		٥٢

$$\bar{س} = \frac{\text{مجـ س ك}}{\text{مجـ س}} = \frac{٦٠٠}{٤} = ١٥$$

$$\text{الانحراف المتوسط} = \frac{\text{مجـ} \left| س - \bar{س} \right| ك}{\text{مجـ ك}} = \frac{٥٢}{٤٠} = ١,٣ \text{ سنة}$$

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المتوسط}}{\bar{س}} \times ١٠٠$$

$$= \frac{١,٣}{١٥} \times ١٠٠ = ٨,٧\%$$

مثال (٣٥):

أوجد الوسط الحسابى والانحراف المعيارى ومعامل الاختـــلاف مـــن التوزيع التالى:

العمر	١٠-	١٥-	٢٠-	٢٥-	٣٠-٣٥	المجموع
ك	٤	١٠	٩	١٢	٥	٤٠

الحـــــل

العمر	ك	س	ح	ح ك	ح٢ ك
١٠-	٤	١٢,٥	١٠-	٤٠-	٤٠٠
١٥-	١٠	١٧,٥	٥-	٥٠-	٢٥٠
٢٠-	٩	٢٢,٥	صفر	صفر	صفر
٢٥	١٢	٢٧,٥	٥+	٦٠+	٣٠٠
٣٠-٣٥	٥	٣٢,٥	١٠+	٥٠+	٥٠٠
المجموع	٤٠			٢٠	١٤٥٠

$$\text{الوسط الحسابى} = \text{الوسط الفرضى} + \frac{\text{مجـ ح ك}}{\text{مجـ ك}}$$

$$= ٢٢,٥ + \frac{٢٠}{٤٠} = ٢٣$$

$$\text{الانحراف المعيارى} = \sqrt{\frac{\text{مجـ ح}^{٢}\text{ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ ح ك}}{\text{مجـ ك}}\right)^{٢}}$$

$$= \sqrt{\frac{١٤٥٠}{٤٠} - \left(\frac{٢٠}{٤٠}\right)^{٢}}$$

$$= \sqrt{٣٦,٢٥ - ٠,٢٥} = \sqrt{٣٦} = ٦$$

$$\text{معامل الاختلاف} = \frac{\text{ع}}{\text{س}} \times ١٠٠ = \frac{٦}{٢٣} \times ١٠٠ = ٢٦\%$$

مثال (٣٦):

الآتى بيان توزيعين تكرارين للأجر الشهرى بالجنيهات وكميات الإنتـاج الأسبوعى لمجموعتين من العمال.

والمطلوب :

إجراء مقارنة بين تشتت الأجور وتشتت كميات الإنتاج للمجموعتين:

فئات الأجر	عدد العمال
-٧	٤
-١٢	١٠
-١٧	٩
-٢٢	١٢
٢٧-٣٢	٥
المجموع	٤٠

فئات الانتاج	عدد العمال
-٢٠	٣
-٢٢	٧
-٢٤	٢١
-٢٦	٥
٢٨-٣٠	٤
المجموع.	٤٠

الحـــل

إذا أخذنا الانحراف المعيارى كمقياس للتشتت.

فئات الأجر	ك	س	ح	ح/	ح/ك	ح/٢ك
٧-	٤	٩,٥	-١٠	-٢	-٨	١٦
١٢-	١٠	١٤,٥	-٥	-١	١٠	١٠
١٧-	٩	١٩,٥	صفر	صفر	صفر	صفر
٢٢-	١٢	٢٤,٥	٥	١	١٢	١٢
٢٧-٣٢	٥	٢٩,٥	١٠	٢	١٠	٢٠
المجموع	٤٠				٤	٥٨

$$\bar{س} = ١٩,٥ + ٥ \times \frac{٤}{٤٠} = ٢٠ \text{ جنيها}$$

$$ع = ط \times \sqrt{\frac{مجـ ح/^٢ ك}{مجـ ك} - \left(\frac{مجـ ح/ ك}{مجـ ك}\right)^٢}$$

$$= ٥ \times \sqrt{\frac{٥٨}{٤٠} - \left(\frac{٤}{٤٠}\right)^٢}$$

$$= ٥ \times \sqrt{١,٤٥ - ٠,٠١} = ٦$$

$$\text{معامل الاختلاف للأجور} = \frac{\text{الانحراف المعيارى}}{\bar{س}} \times ١٠٠$$

$$= \frac{٦}{٢٠} \times ١٠٠ = ٣٠\%$$

وتجرى نفس الخطوات السابقة بالنسبة للإنتاج .

فئات الأجر	ك	س	ح	ح/	ح/ ك	ح/٢ ك
٢٠-	٣	٢١	-٤	-٢	-٦	١٢
٢٢-	٧	٢٣	-٢	-١	-٧	٧
٢٤-	٢١	٢٥	صفر	صفر	صفر	صفر
٢٦-	٥	٢٧	٢	١	٥	٥
٢٨-٣٠	٤	٢٩	٤	٢	٨	١٦
المجموع	٤٠				صفر	٤٠

$$\bar{س} = ٢٥ + ٢ \times \frac{صفر}{٤٠} = ٢٥$$

$$ع = ط \times \sqrt{\frac{مجـ ح/^{٢} ك}{مجـ ك} - \left(\frac{مجـ ح/ ك}{مجـ ك}\right)^{٢}}$$

$$= ٢ \times \frac{٤٠}{٤٠} = ٢$$

$$\therefore \text{معامل الاختلاف} = \frac{ع}{\bar{س}} \times ١٠٠$$

$$= \frac{٢}{٢٥} \times ١٠٠ = ٨ \%$$

وبمقارنة معامل الاختلاف للاجور بمعامل الاختلاف بالإنتاج يتضــــح أن الأجور أكثر تشتتا من الإنتاج.

مثال (٣٧):

إختيرت عينة عشوائية من خمسة طلبة من أحد المعاهد الفنية فكانت أعمارهم ٢٠ ، ٢٨ ، ٢٧ ، ١٩ ، ٢٠ بينما أوضحت عينة عشوائية من أربعة طلبه من كلية التجارة أن أعمارهم كما يلي: ٢٢ ، ١٨ ، ١٩ ، ٢٤ والمطلوب: استخدام المدى لمقارنة التشتت فى العمر بين طلبة المعهد وطلبة كلية التجارة.

الحل

المدى لطلبة المعهد = أكبر قيمة − أقل قيمة = ٢٨ − ٢٠ = ٨

$$\text{معامل الاختلاف} = \frac{٢٨ - ٢٠}{٢٨ + ٢٠} \times ١٠٠ = ١٧\%$$

المدى لطلبة كلية التجارة = ٢٤ − ١٨ = ٦

$$\text{معامل الاختلاف} = \frac{٢٤ - ١٨}{٢٤ + ١٨} \times ١٠٠ = ١٤,٣\%$$

مثال (٣٨):

من التوزيع التكرارى التالى المطلوب حساب الإنحراف المتوسط من الوسط الحسابى .

الفئات	٩ −	١٣ −	١٧ −	٢١ −	٢٥ − ٢٩
التكرار	٥	١٥	٢٥	٢٥	١٠

الحل

لحساب الانحراف المتوسط عن الوسط الحسابى لابد أن نقوم أولا بحساب الوسط الحسابى أولا.

فئات	ك	س	ح	ح/	ح/ك	|س - س̄|	|س - س̄| ك
٩-	٥	١١	-٨	-٢	-١٠	٩	٤٥
١٣-	١٥	١٥	-٤	-١	-١٥	٥	٧٥
١٧-	٢٥	١٩	صفر	صفر	صفر	١	٢٥
٢١-	٢٥	٢٣	٤	١	٢٥	٣	٧٥
٢٥-٢٩	١٠	٢٧	٨	٢	٢٠	٧	٧٠
المجموع	٨٠				٢٠		٢٩٠

الوسط الحسابى = الوسط الفرضى + (مقدار ثابت × مجـ ح/ ك) / مجـ ك

= ١٩ + (٤ × ٢٠) / ٨٠

= ١٩ + ١٠ = ٢٠

الانحراف المتوسط = مجـ |س - س̄| ك / مجـ ك

= ٢٩٠ / ٨٠ = ٣,٦٢٥

مثال (٣٩):

من الجدول التكرارى التالى أوجد الوسط الحسابى والأنحراف المعيـــارى ومعامل الاختلاف واذكر فائدة معامل الاختلاف.

فئات السن	-٥	-١٠	-١٥	-٢٠	٢٥-٣٠	المجموع
التكرارى	٨	٢٠	١٨	٢٤	١٠	٨٠

لحساب الأنحراف المعيارى نتبع نفس الخطوات التى اتبعناها سابقا عند حساب الوسط الحسابى.

فئات السن	ك	س	ح	ح/	ح/ك	ح/٢ك
-٥	٨	٧,٥	-١٠	-٢	-١٦	٣٢
-١٠	٢٠	١٢,٥	-٥	-١	-٢٠	٢٠
-١٥	١٨	١٧,٥	صفر	صفر	صفر	صفر
-٢٠	٢٤	٢٢,٥	+٥	+١	+٢٤	٢٤
٢٥-٣٠	١٠	٢٧,٥	+١٠	+٢	+٢٠	٤٠
المجموع	٨٠				+٨	١١٦

$$\text{الوسط الحسابى} = ٥.١٧ + \frac{٥ \times ٨}{٨٠} = ١٨ \text{ سنه}$$

$$\text{الانحراف المعيارى} = \text{ط} \times \sqrt{\frac{\text{مجـ ح}^{/٢}\text{ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ ح}^{/}\text{ك}}{\text{مجـ ك}}\right)^{٢}}$$

$$ع = ٥ \times \sqrt{\frac{١١٦}{٨٠} - \left(\frac{٨}{٨٠}\right)^{٢}}$$

$$= ٥ \times \sqrt{١,٤٥ - ٠,٠١} = \sqrt{١,٤٤} \times ٥$$

$$= ٥ \times ١,٢ = ٦ \text{ سنوات}$$

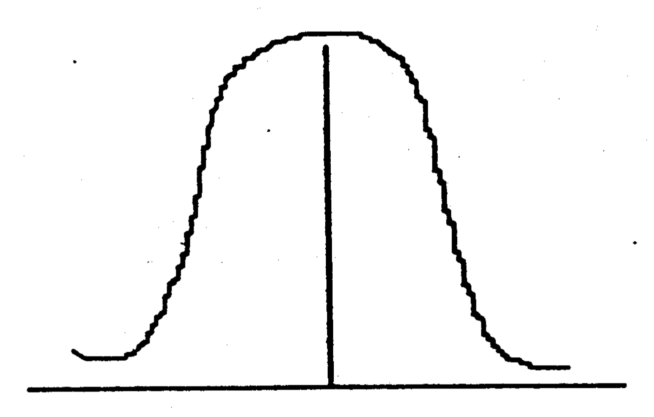
ونلاحظ أن الأنحراف المعيارى يعتبر مقياس مطلق للتشتت بمعنى أن له تمييز (جنيه أو سنة أو قرش .. ألخ) فاذا أردنا المقارنة بين تشتت توزيعين تكرارين فلا يجب المقارنة على أساس الأنحراف المعيارى لكل منهما وإنما لابد من حساب معامل الاختلاف حيث.

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعيارى}}{\text{الوسط الحسابى}} \times ١٠٠$$

$$= \frac{٦}{١٨} \times ١٠٠ = ٣٣,٣\%$$

رابعا: الالتـــواء:

تمثل دراسة الالتواء للتوزيع التكرارى الخطوة الرابعـــة فـــى التحليـل الاحصائى للتوزيع. فبعد أن درسنا متوسط التوزيع وناقشـــنا درجـة تركـز مفردات الظاهرة حول المتوسط (التشتت) ثم أردنا التعبير عن التشتت المطلـق فى صورة نسبية (معامل الاختلاف) يهمنا الأن دراسة درجة التواء المنحنـــى التكرارى للتوزيع ويطلق على المنحنى الموضح فى الشكل التالى بأنه منحنـى متماثل.



شكل (٩)

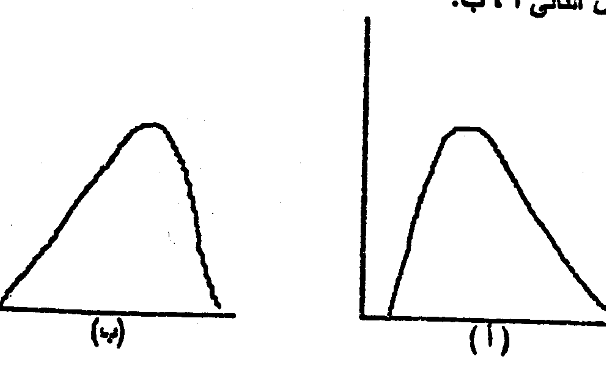
واذا أسقطنا عمودا من قمة المنحنى لوجدنا أن المنحنى قد انشـــطر الـــى نصفين متماثلين ومن هنا يطلق على المنحنى بأنه منحنى متمـــاثل. ويعتـــبر منحنى التوزيع الطبيعى أشهر المنحنيات المتماثلة بل إن البعض قد يقرن صفة الطبيعى مع صفة التماثل.

ونجد فى المنحنى المتماثل أن قيمة الوسط الحسابى للتوزيع تساوى قيمة الوسيط تساوى قيمة المناول وذلك فى النقطة التى تقابل قمة المنحنى.

وقد لاتجد المنحنى متماثلا بالضبط ولكنه قريبا من التماثل وفى هذه الحالة توجد علاقة تقريبية تحكم كل من الوسط الحسابى والوسيط والمنوال هى:

الوسط الحسابى - المنوال = ٣ (الوسط الحسابى - الوسيط) ....(٢٨)

ولكن يحدث كثيرا أن نجد أن نصفى المنحنى غير متماثلين مثل ماهو موضح فى الشكل التالى أ ، ب:



شكل (١٠)

ففى الشكل (أ) نلاحظ أن قمة المنحنى تتجه الى اليسار كما أن مفردات التوزيع تأخذ قيما تتدرج ببطىء أولا حتى تصل الى أعلى قيمة فى المنحنى ثم تبدأ قيم التوزيع فى التناقص بسرعة حتى تصل الى نهاية المنحنى ويطلق على هذا المنحنى أنه منحنى موجب الالتواء. ونجد فى المنحنى موجب الالتواء أن الوسط الحسابى دائما أكبر من المنوال، كما أن الوسيط يقع بينهما.

بينما نلاحظ فى الشكل (ب) أن قمة المنحنى تتجه نحو اليمين كما أن مفردات التوزيع تأخذ قيما تتدرج بسرعة فى الزيادة حتى تصل الى قمة المنحنى ثم تبدأ فى التناقص ببطىء حتى نهاية المنحنى، ونطلق على هذا النوع من المنحنيات بأنه منحنى سالب الالتواء حيث نجد أن قيمة المنوال أكبر من قيمة الوسط الحسابى وأيضا نجد أن قيمة الوسيط تقع بين المنوال والوسط الحسابى.

ولكى تكتمل الصورة للتوزيع التكرارى فانه يجب تحديد طبيعة ودرجة الالتواء للمنحنى التكرارى.

ويعتبر قياس الالتواء بإستخدام العزوم أفضل الطرق لتحديد الالتواء للتوزيع التكرارى، وسوف نناقش ذلك فى الباب السابع، ولهذا سوف نستخدم العلاقات الثلاث التالية فى دراسة الالتواء للمنحنى التكرارى وهى تعتمد على ماسبق دراسته من المقاييس الاحصائية:

$$\text{معامل الالتواء الأول} = \frac{\text{الوسط الحسابى} - \text{المنوال}}{\text{الانحراف المعيارى}} \quad ....(٢٩)$$

$$\text{معامل الالتواء الثانى} = \frac{٣\,(\text{الوسط الحسابى} - \text{الوسيط})}{\text{الانحراف المعيارى}} \quad ....(٣٠)$$

ويطلق على هذه المعاملات اسم معاملات إلتواء بيرسون ويمكن اثبات أن قيمتها تنحصر بين + ٣ ، - ٣

كما يمكن استخدام الوسيط والربيع الأول والربيع الثالث فى قياس الالتواء حيث يطلق عليه معامل التواء بولى.

$$\text{معامل الالتواء} = \frac{(\text{الربيع الثالث} + \text{الربيع الأول} - ٢ \times \text{الوسيط})}{(\text{الربيع الثالث} - \text{الربيع الأول})} \quad ....(٣١)$$

ويجب أن نلاحظ أن إشارة القيمة التى نحصل عليها لمعامل الالتواء توضح طبيعة الالتواء، فالاشارة الموجبة لمعامل الالتواء توضح أن المنحنى موجب الالتواء، كما أن الاشارة السالبة لمعامل الالتواء توضح أن المنحنى سالب الالتواء. أما اذا كان معامل الالتواء يساوى الصفر فهذا يعنى تماثل منحنى التوزيع التكرارى.

أيضا يجب أن نلاحظ أن قيمة معامل الالتواء تعتبر قيمة نسبية بمعنى أن قيمة معامل الالتواء لاتأخذ وحدة القياس الخاصة بالظاهرة، وقد يرغب البعض فى التأكيد على الصورة النسبية لمعامل الالتواء بجعل قيمته نسبة مئوية بالضرب × ١٠٠ ولكننا لانرى مبررا لذلك فيكفى أن نؤكد أن معامل الالتواء لايأخذ أى وحدة قياس خاصة بالظاهرة.

مثال (٤٠):

أوجد معامل الالتواء للتوزيع التكرارى للطول الموضح فى مثال (٣١)

الحل

نلاحظ من مثال (٣١) أننا حصلنا على القيم التالية:

الوسيط = ١٧٤,٤ سم

الربيع الأول = ١٦٨ سم

الربيع الثالث = ١٨٠ سم

وباستخدام المعادلة (٣١) نجد أن:

$$\text{معامل الالتواء} = \frac{(\text{الربيع الثالث} + \text{الربيع الأول} - ٢ \times \text{الوسيط})}{(\text{الربيع الثالث} - \text{الربيع الأول})}$$

$$= \frac{(١٨٠ + ١٦٨ - ٢ \times ١٧٤,٤)}{(١٨٠ - ١٦٨)}$$

$$= - ٠,٨ \div ١٢ = - ٠,٠٧$$

ويتضح من ذلك أن المنحنى سالب الالتواء.

مثال (٤١):

إدرس الالتواء لتوزيع العمر الموضح فى مثال (٣٣)

الحـــل

نلاحظ من مثال (٣٣) أننا حصلنا على:

الوسط الحسابى = ٣١,٦ سنة

الانحراف المعيارى = ٩,٥ سنة

ومن ثم يمكننا إيجاد المنوال حتى يمكن إستخدام المعادلـــة (٢,٩) لإيجـــاد الالتواء.

فاذا اخترنا طريقة الرافعة لايجاد المنوال فإن:

$$\text{المنوال} = ٢٦ + ٨ \times \frac{٢٠}{٢٠ + ١٦}$$

$$= ٢٦ + ٤,٤ = ٣٠,٤ \text{ سنة}$$

$$\text{معامل الالتواء} = \frac{\text{(الوسط الحسابى - المنوال)}}{\text{الانحراف المعيارى}}$$

$$= \frac{٣١,٦ - ٣٠,٤}{٩,٥} = + ٠,١٢$$

مما يوضح أن المنحنى موجب الالتواء

مثال (٤٢):

إستخدام الوسيط فى قياس الالتواء للتوزيع التكرارى فى مثال (٣٣)

الحـــل

يجب إيجاد الوسيط من التوزيع التكرارى للعمـــر. وبتكويـــن المنحنـــى المتجمع الصاعد.

أقل من حدود الفئات	التكرار المتجمع الصاعد
أقل من ١٠	صفر
أقل من ١٨	٤
أقل من ٢٦	٢٠
أقل من ٣٤	٥٠
أقل من ٤٢	٧٠
أقل من ٥٠	٨٠

$$\text{رتبة الوسيط} = \frac{\text{مجـ ك}}{٢}$$

$$ك_{٢} = \frac{٨٠}{٢} = ٤٠$$

وبالبحث عن رتبة الوسيط بين التكرارات المتجمعة نجد أنها تقع بين ٢٠، ٥٠
ومن ثم يكون تكرار الفئة الوسيطية = ٣٠ = ك
كما أن:

$$\text{الوسيط} = \text{بداية فئة الوسيط} + \frac{ط\,(ك_2 - ك_1)}{ك_3 - ك_1}$$

$$= ٢٦ + \frac{٨\,(٤٠ - ٢٠)}{٥٠ - ٢٠} = ٣١٫٣ \text{ سنة}$$

وحيث أن :

الوسط الحسابى = ٣١٫٣ سنة

الانحراف المعيارى = ٩٫٥ سنة

ومن العلاقة (٣٠) فإن:

$$\text{معامل الالتواء} = \frac{٣\,(\text{الوسط الحسابى} - \text{الوسيط})}{\text{الانحراف المعيارى}}$$

$$= \frac{٣\,(٣١٫٦ - ٣١٫٣)}{٩٫٥}$$

$$= + ٠٫٠٩$$

مما يعنى أن المنحنى موجب الالتواء

مثال (٤٣):

فى عينة مكونة من طلبة جامعة القاهرة أو ضحت أن متوسط الدخل الشهرى ٢٥٠ جنيها والانحراف المعيارى ٤٠ جنيها وفى عينة أخرى من طلبة جامعة عين شمس أوضحت أن:

الدخل الشهرى	٧٠-	٨٠-	٩٠-	١٠٠-	١١٠-	١٢٠-١٣٠
عدد الأشخاص	٧	٩	٢٨	٣٠	٢٠	٦

والمطلوب:

(أ) فى أى جامعة من الجامعتين يعتبر الدخل أكثر تشتتا.

(ب) إحسب معامل الألتواء للدخل لطلبة جامعة عين شمس.

الحـــل

لمعرفة أى الجامعتين يعتبر الدخل (أكثر تغيرا لابد من حساب أحد مقاييس التغير أو التشتت من الجدول الثانى وحيث أن قد أعطى لنا الانحراف المعيارى للعينه الأولى فلابد من حساب الانحراف المعيارى للعينه الثانية – ولكن لايجب مقارنة التغير على أساس الانحراف المعيارى لكل منهما ولابد من حساب قابل الاختلاف لكل توزيع حتى يمكن استخدامه فى المقارنة بينهما.

الدخل الشهرى	ك	س	ح	ح/	ح/ك	ح/٢ك
٧٠-	٧	٧٥	-٣٠	-٣	-٢١	٦٣
٨٠-	٩	٨٥	-٢٠	-٢	-١٨	٣٦
٩٠-	٢٨	٩٥	-١٠	-١	-٢٨	٢٨
١٠٠-	٣٠	١٠٥	صفر	صفر	صفر	صفر
١١٠-	٢٠	١١٥	١٠	١	٢٠	٢٠
١٢٠-١٣٠	٦	١٢٥	٢٠	٢	١٢	٢٤
المجموع	١٠٠				-٣٥	١٧١

$$\overline{\text{س}} = ١٠٥ + \frac{١٠ \times -٣٥}{١٠٠} = ١٠١,٥ \text{ جنيه}$$

$$\text{ع} = \text{ط} \times \sqrt{\frac{\text{مجـ ح}^{/٢}\text{ك}}{\text{مجـ ك}} - \left(\frac{\text{مجـ ح}^{/}\text{ك}}{\text{مجـ ك}}\right)^{٢}}$$

$$\text{ع} = ١٠ \times \sqrt{\frac{١٧١}{١٠٠} - \left(\frac{٣٥}{١٠٠}\right)^{٢}}$$

$$= ١٠ \times \sqrt{١,٧١ - ٠,١٢٢٥} = ١٢,٦ \text{ جنيها}$$

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعيارى}}{\text{الوسط الحسابى}} \times ١٠٠$$

$$\therefore \text{معامل الاختلاف لجامعة القاهرة} = \frac{٤٠}{٢٥٠} \times ١٠٠ = ١٦ \%$$

$$\therefore \text{معامل الاختلاف لجامعة عين شمس} = \frac{١٢,٦}{١٠١,٥} \times ١٠٠ = ١٢,٤\%$$

من هذا يتضح أن دخل الطلبة فى جامعة القاهرة أكثر تغيرا من دخل الطلبة فى جامعة عين شمس.

بالنسبة لمعامل الالتواء فهناك طرق عديدة لإيجاده ولكن حتى يمكن الاستفاده بالعمليات الحسابية السابقة سنستخدم الصورة التالية لمعامل الالتواء.

$$\text{معامل الالتواء} = \frac{\text{الوسط الحسابى} - \text{المنوال}}{\text{الانحراف المعيارى}}$$

ولذا لابد من حساب المنوال بأى طريقة من الطرق

المنوال بطريقة الرافعه = ١٠٠ + $\frac{١٠ \times ٢٠}{٢٨ + ٢٠}$

= ١٠٠ + ٢.٤ = ١٠٤,٢

معامل الالتواء = $\frac{١٠١,٥ - ١٠٤,٢}{١٢,٦}$

= $\frac{- ٢,٧}{١٢,٦}$ = - ٠,٢١

وبالتالى فهو سالب الالتواء

مثال (٤٤):

أوجد الوسيط والربيع الأول والربيع الثالث بالرســـم وأوجـد معــامل الالتواء من التوزيع التالى:

العمر	١٠-	١٣-	١٦-	١٩-	٢٢-٢٥	المجموع
ك	٥	١٥	٣٠	٢٠	١٠	٨٠

الحـــل

أقل من حدود الفئات	التكرار المتجمع الصاعد
أقل من ١٠	صفر
أقل من ١٣	٥
أقل من ١٦	٢٠
أقل من ١٩	٥٠
أقل من ٢٢	٧٠
أقل من ٢٥	٨٠

رتبة الربيع الأول = $\frac{\text{٨٠}}{\text{٤}}$ = ٢٠

رتبة الربيع الثالث = $\frac{\text{٣} \times \text{٨٠}}{\text{٤}}$ = ٦٠

رتبة الوســـيط = $\frac{\text{٨٠}}{\text{٢}}$ = ٤٠

ويتضح من الرسم أن: قيمة $\text{ب}_{١}$ = ١٦

وقيمة $\text{ب}_{٢}$ = ١٨

وقيمة $\text{ب}_{٣}$ = ٢٠,٥

$$\text{معامل الالتواء} = \frac{\text{ب}_{٣} + \text{ب}_{١} - \text{٢ ب}_{٢}}{\text{ب}_{٣} + \text{ب}_{١}}$$

$$\text{معامل الالتواء} = \frac{\text{٢٠,٥} + \text{١٦} - \text{٢} \times \text{١٨}}{\text{٢٠,٥} - \text{١٦}} = \frac{+\text{٠,٥}}{\text{٤,٥}} = \text{١١ر}$$

تمارين

١) يمثل التوزيع التالى التوزيع التكرارى لعينة من ١٠٠ طالب بكلية التجارة على فئات الذكاء المختلفة :

الذكاء	-٥٠	-٦٠	-٧٠	-٨٠	٩٠-١٠٠	المجموع
ك	١٥	٢٥	٤٠	١٦	٤	١٠٠

والمطلوب : إيجاد كل من الوسط الحسابى والوسيط والمنوال .

٢) أوجد الوسيط بالحساب وبالرسم للتوزيع التالى:

الدخل	-٢٠٠	-٢٣٠	-٢٥٠	-٣٠٠	-٣٥٠	٤٠٠-٤٢٠	المجموع
ك	٦	٢٤	٥٠	٤٠	١٥	٥	١٤٠

٣) أوجد المنوال بالحساب وبالرسم بطريقة فروق التكرارات للتوزيع التالى:

الوزن	-٥٠	-٥٥	-٦٥	-٧٠	٨٠-٩٠	المجموع
ك	١٠	٤٠	٤٠	٦٠	٢٠	١٧٠

٤) إذا كانت أعمار سبعة من أصدقائك هى:

٢٢ ، ٢٠ ، ٢٣ ، ٢٥ ، ٢٧ ، ١٩ ، ٢٥ سنة

فما هى قيمة كل من الوسط الحسابى، الوسيط ، المنوال للعمر.

٥) يوضح الجدول التالى التوزيع التكرارى للدخل الشهرى لمجموعة مـــن العاملين بالجامعة.

الدخل	٢٠٠-	٢٤٠-	٢٦٠-	٢٨٠-	٣٠٠-	٣٢٠-٣٤٠	المجموع
ك	٧	٢٣	٤٠	٢٠	٨	٢	١٠٠

والمطلوب إيجاد كل من:

أ - الوسط الحسابى للدخل.

ب - الوسيط بالحساب وبالرسم.

جـ- المنوال بالحساب وبالرسم.

٦) أوجد المنوال بالحساب وبالرسم بطريقة فروق التكرارات من التوزيع التالى:

الأجر	١٠٠-	١٢٠-	١٥٠-	٢٠٠-	٢٢٠-	٢٣٠-٢٥٠
ك	٤٠	٩٠	٤٠٠	٢٨٠	٩٠	٥٠

٧) أوضحت عينة عشوائية من سبعة طلاب أن الدخل الشهرى كما يلى بالجنيه:

٦٠، ٤٨، ٦٤، ٤٤، ٢٤، ٧٥، ٦٣

والمطلوب : (أ) إيجاد الانحراف المعيارى ومعامل الاختلاف.

(ب) إيجاد المدى ومعامل الاختلاف.

(جـ) الانحراف المتوسط ومعامل الاختلاف.

٨) يوضح الجدول التالى التوزيع التكرارى لعينة عشوائية من ٥٠٠ طـــالب من كلية التجارة موزعة طبقا للذكاء.

الذكـــــاء	٤٠-	٥٠-	٦٠-	٦٥-	٧٥-	٨٠-٨٥
عدد الطلاب	٢٨	٤٢	١٤٠	١٩٠	٨٠	٢٠

والمطلوب: إيجاد الوسيط والربيعين بالحساب وبالرسم وإيجاد معامل الالتــواء (معامل بولى).

٩) أختيرت عينة عشوائية من خمسة طلاب من بين طلبة الفرقتين الأولـــى والثانية بأحد المعاهد وتم إجراء اختبار للذكاء لهم فكانت نتائج العينـــات كما يلى:-

نتائج العينة من الفرقة الأولى : ٦٠ ، ٧٨ ، ٨٢ ، ٦٠ ، ٧٥

نتائج العينة من الفرقة الثانية : ٩٣ ، ٩٧ ، ٩٠ ، ٨٥ ، ٩٥

والمطلوب : (أ) إيجاد الوسط الحسابى والمدى والانحراف المعيارى.

(ب) المقارنة بين تشتت الفرقتين بطريقتين مختلفتين.

١٠) أوجد المنوال بالحساب وبالرسم من الجدول التالى:

الدخل الشهرى بالجنيه	٢٥٠-	٢٨٠-	٣١٠-	٣٤٠-	٣٧٠-٤٠٠
ك	٧	٢٥	٥٠	١٥	٣

١١) أوضحت إحدى الدراسات أن الوسط الحسابى لدخل الاسرة فى مدينة الاسكندرية هو ٣٢٠ جنيها شهريا والانحراف المعيارى للدخل ٢٥ جنيها بينما كان التوزيع التكرارى لعينة من القاهرة كما يلي:-

فئات الدخل	٣١٠-	٣٢٠-	٣٣٠-	٣٤٠-	٣٥٠-٣٦٠
عدد الأسر	٥	١٥	٤٠	٣٠	١٠

والمطلوب: (أ) المقارنة بين تشتت الدخل فى المدينتين.

(ب) قياس الالتواء للدخل من عينة أسر القاهرة.

١٢) يوضح التوزيع التكرارى توزيع عينة عشوائية من ٢٧٠ طالب على فئات الوزن المختلفة

فئات الوزن	٦٤-	٧٠-	٨٠-	٩٠-	٩٥-٩٨
عدد الطلبة	٢٤	٩٠	١٢٠	٣٠	٦

أوجد : أ) المنوال بالحساب وبالرسم.

ب) المدى.

١٣) أوجد الانحراف المتوسط عن الوسط الحسابى من التوزيع التالى ثم أوجد معامل الاختلاف.

المكافآت السنوية	٣٠٠-	٣٥٠-	٤٠٠-	٤٥٠-	٥٠٠-٥٥٠
عدد العاملين	٨	١٢	٤٠	٢٥	٥

١٤) تبين من التوزيع التكرارى للمكافآت السنوية فـــى إحـــدى المؤسســـات التجارية الخاصة أن : الوسط الحسابى للمكافآت السنوية = ٥٠٠ جنيه

والانحراف المعيارى = ٧٥ جنيها

المنوال = ٤٢٠ جنيها

والمطلوب: إيجاد معامل الالتواء مع توضيح مايدل عليه.

١٥) أوجد معامل الالتواء للتوزيع التالى:

الأجر	-٥٠	-٦٠	-٧٠	-٨٠	٩٠-١٠٠	المجموع
ك	٦٥	١٤	٤٠	٢٥	١٥	١٠٠

الباب الرابع

الارتباط الخطى

الارتباط الخطى البسيط

تمهيـد :

من الاهداف الهامة فى كثير من الدراسات الاقتصادية أو الإدارية أو الطبيعية أو الإجتماعية . . . إلخ ، هو دراسة العلاقة بين ظاهرتين أو أكثر بغرض الوقوف على طبيعة هذه العلاقة ، سواء من حيث معرفة درجة قوتها أو ضعفها من ناحية ، أو من حيث تعيين اتجاهها من ناحية أخرى ، ويعتبر الارتباط من الأدوات الإحصائية الهامة التى تفيدنا فى هذا المجال .

هذا والأمر لا يقتصر على دراسة العلاقـة بين ظاهـرتين أو أكـثر على دراسـة الارتباط بينـهما ، بل كثيراً ما يتطلب الأمر التنـبؤ لما سـوف تكون عليه ظـاهرة ما فى المسـتقبل ، على ضـوء القـيم التى تأخذها ظاهـرة أو عدة ظواهر أخـرى تكـون لها اعـتماداً كبيراً عليها . وهذا التنـبؤ غالباً ما يكون له فائدة كبـيرة فى اتخـاذ قرارات معينة ، كما هو الحال بالنسـبة للتنبؤ بالطلب على سلعة معينه أو بالمبيـعات أو الأربـاح أو العـمالة المسـتقبلة . إلا أن درجة الدقـة المتوقـعة من اسـتخدام النـماذج الرياضية Mathematical Models فى التنـبؤ والتى قد تكـون معادلة أو مجمـوعة من المعـادلات ، تتـوقف على عدة عـوامل كثيرة منها :

أولاً : درجة الدقة فى البيانات المتاحة والمستخدمة فى النموذج .

ثانياً : مدى احتواء النمـوذج على جمـيع المتغـيرات الـتى تؤثر على المشـكلة موضـوع الدراسـة .

ثالثاً : مـدى قرب النمـوذج للواقـع العمـلى .

هذا ويعتبر تحليل الانحدار regression analysis من الأدوات الاحصائية الهامة التى تفيدنا فى مجال التنبؤ .

وفى هذا الباب سوف نركز اهتمامنا على دراسة الارتباط بين متغيرين

وتجدر الاشارة أنه فى دراسة الارتباط بين متغيرين وليكن س ، ص فإنه ليس من الضرورى أن يكون أحداهما معتمداً على الآخر فى تغيره ، بمعنى أنه ليس من المهم التمييز بين أيهما المتغير التابع وأيهما المتغير المستقل ، لأن الارتباط بين المتغيرين كما سيتبين لنا فيما بعد ، يوضح أن المتغيرين موضوع الدراسة يسلكان سلوكاً معيناً ، حيث يكون التغير بالزيادة أو بالنقص فى أحداهما يكون مصحوباً بتغير فى الآخر سواء أكان هذا التغير فى نفس الاتجاه أو فى الاتجاه المضاد . ويترتب على ذلك أن قيم المتغيرين س ، ص تعتبر أزواجاً عشوائية من المشاهدات التى لا يتحكم الباحث فى أى منها . فمثلا إذا اخترنا عاملا من مصنع معين وأخذنا مدة خدمته بالسنوات وأجره بالجنيهات ، فمن الواضح أن الباحث يسجل هنا ما يشاهده دون أن يتحكم من جانبه فى أيهما ، وبذلك فإن مدة الخدمة والأجر قد تركتا لأن تأخذا أى قيمة ممكنة . وبالمثل إذا اخترنا طالباً من سنة معينة وأخذنا درجاته فى مادتى الرياضة البحتة والاقتصاد ، فمن الواضح أن الباحث هنا يسجل ما حصل عليه هذا الطالب من درجات فى هاتين المادتين ، وبذلك فإن درجات الرياضة البحتة ودرجات الاقتصاد قد تركتا لأن تأخذا أى قيمة ممكنة وهكذا . . .

شكل الانتشار :

فى تحليل العلاقة بين متغيرين نجد أنه من المفضل عادة عرض أزواج القيم المتناظره لهذين المتغيرين فى شكل بيانى : فإذا ما أخذنا المحور الأفقى لتمثيل قيم أحد المتغيرين وليكن س والمحور الرأسى لتمثيل قيم المتغير الآخر وليكن ص فإن كل زوج من أزواج القيم المتناظرة لهذين المتغيرين ولتكن :

$$(س_١ ، ص_١) ، (س_٢ ، ص_٢) ، \ldots\ldots ، (س_ن ، ص_ن)$$

يمكن تمثيلها بيانياً بنقطة ، وتكون جميع النقط ما يعرف بشكل الانتشار وهذا الشكل يعطينا صورة مرئية عن طبيعة العلاقة بين هذين المتغيرين ، فضلا عن اقتراح النموذج الذى يلائم البيانات موضوع الدراسة .

وشكل الانتشار قد يأخذ إحدى الصور الآتية :

الأولى : من الشكل المقابل يتبين أن نقط شكل الانتشار تنتشر حول خط مستقيم يتجه من أدنى اليسار إلى أعلى اليمين . وهذا يوضح أن المتغيرين س ، ص مرتبطان بعلاقة طردية أو موجبة . وهذا يعنى أن كلّ زيادة فى المتغير س يصاحبها زيادة فى المتغير الآخر ص

ص

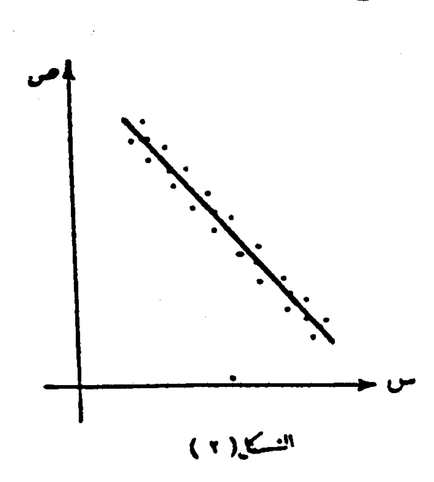
س

الشكل (١)

وبالعكس كل نقص فى المتغير س يصاحبها نقص فى المتغير ص . وبعبارة أخرى فإن التغير فى كل من هذين المتغيرين سواء بالزيادة أو بالنقص يسيران فى نفس الاتجاه .

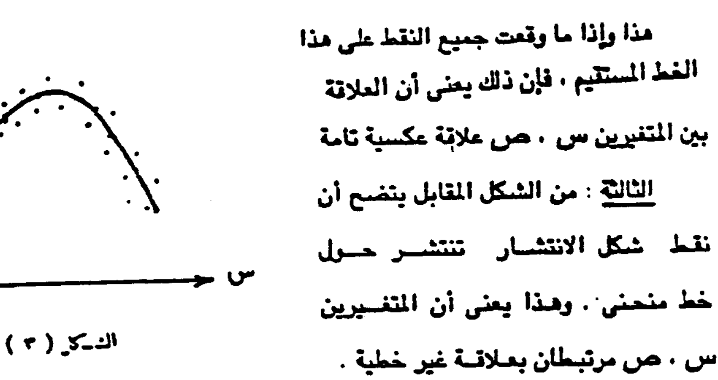
هذا وإذا ما وقعت جميع النقط الممثلة لأزواج القيم على الخط المستقيم ، فإن ذلك يعنى أن هناك علاقة طردية تامة بين المتغيرين س ، ص .

الثانية : من الشكل المقابل يتبين أن نقط شكل الانتشار تنتشر حول خط مستقيم ، ولكن بصورة مخالفة للحالة السابقة . وهذا يوضح أن المتغيرين س ، ص مرتبطان بعلاقة عكسية أو سالبة . وهذا يعنى أن كل زيادة فى المتغير س يصاحبها نقص فى المتغير الآخر ص . وبالعكس ، كل نقص فى المتغير س يصاحبه زيادة فى المتغير ص .



الشكل (٢)

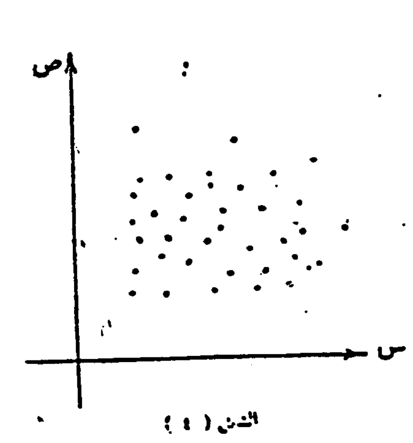
وبعباره أخرى فإن التغير فى كل من هذين المتغيرين يسيران فى اتجاهين متضادين .



الشكل (٣)

هذا وإذا ما وقعت جميع النقط على هذا الخط المستقيم ، فإن ذلك يعنى أن العلاقة بين المتغيرين س ، ص علاقة عكسية تامة

الثالثة : من الشكل المقابل يتضح أن نقط شكل الانتشار تنتشر حول خط منحنى . وهذا يعنى أن المتغيرين س ، ص مرتبطان بعلاقة غير خطية .



الشكل (٤)

الرابعة : من الشكل المقابل يتضح ان نقط شكل الانتشار لا تنتشر حول خط مستقيم أو منحنى بل نجدها مبعثرة فى جميع أنحاء الشكل البيانى بشكل غير منتظم ، وبما يعنى عدم وجود علاقة بين المتغيرين س ، ص موضوع الدراسة . وبعبارة أخرى فإن هذين المتغيرين يعتبران مستقلان .

هذا والعلاقة الموضحة بالشكلين (١) ، (٢) والتى تعرض لنا النقط الممثلة لأزواج قيم الظاهرتين س ، ص متجمعة حول مستقيم يطلق عليها بالارتباط الخطى linear correlation أما العلاقة الموضحة بالشكل (٣) والتى تعرض تجمع نقط شكل الانتشار حول خط منحنى فيطلق عليها بالارتباط غير المستقيم nonlinear correlation أما العلاقة الموضحة بالشكل (٤) والتى تعرض نقط شكل الانتشار بشكل غير منتظم فتعنى أن الارتباط بين المتغيرين منعدم .

وعلى ضوء ما تقدم ، فإنه يتضح لنا أن شكل الانتشار يعطى فكرة تقريبية عن طبيعة العلاقة بين المتغيرين . فوقوع النقط الممثلة لأز واج القيم حول خط مستقيم أو منحنى ، يعنى وجود علاقة قوية بين المتغيرين . وكلما تباعدت هذه النقط كلما دل ذلك على ضعف هذه العلاقة . أما إذا كانت النقط منتشرة بشكل غير منتظم فإن ذلك يعنى عدم وجود ثمة علاقة بين المتغيرين .

هذا ولا يكتفى بدراسة العلاقة بين متغيرين على الصورة البيانية ، بل أن الأمر يتطلب وجود مقاييس لقياس درجة العلاقة بين المتغيرين س ، ص بطريقة كمية من ناحية ، ولتعيين اتجاهها من ناحية أخرى .

هذا وإذا ما اقتصرت دراسة الارتباط على العلاقة بين متغيرين إثنين فقط ، فإن ذلك يطلق عليه بالارتباط البسيط . أما إذا كانت العلاقة تتضمن أكثر من متغيرين ، فإنه يطلق عليه بالارتباط المتعدد أو الجزئى . وعلى أية حال فإننا سوف نقتصر فى دراستنا للارتباط على الارتباط الخطى البسيط .

معامل الارتباط الخطى البسيط

يستخدم معامل بيرسون للارتباط في حساب الارتباط الخطى البسيط للمتغيرين س ، ص سواء أكانت البيانات غير مبوبة أو مبوبة فى صورة جدول توزيع تكرارى مزدوج ، وقيمة هذا المعامل تتراوح بين + ١ ، - ١ . هذا ويكون :

أولاً : معامل الارتباط يساوى + ١ :

إذا كانت هناك علاقة طردية تامة بين الظاهرتين س ، ص .

وفى هذه الحالة نجد أن جميع نقط شكل الانتشار تقع على خط مستقيم (ممتد من أدنى اليسار إلى أعلى اليمين) دون أى انحراف عنه . ويترتب على ذلك أن قيم الظاهرة س الكبيرة تكون مصاحبة لقيم الظاهرة ص الكبيرة ، وقيم الظاهرة س الصغيرة تكون مصاحبة لقيم الظاهرة ص الصغيرة . وهذا يعنى أن أكبر قيمة للظاهرة س تناظرها أكبر قيمة للظاهرة ص ، والقيمة الثانية فى الكبر للظاهرة س تناظرها القيمة الثانية فى الكبر للظاهرة ص ، وهكذا بالاستمرار فى التدريج نصل فى النهاية إلى أن تكون أصغر قيمة للظاهرة س تناظرها أصغر قيمة للظاهرة ص .

ومن أمثلة الارتباط الطردى التام ، العلاقة بين طول ضلع المربع ومساحته والعلاقة بين وزن سلعة معينة وسعرها .

ثانياً : معامل الارتباط يساوى - ١ :

إذا كانت هناك علاقة عكسية تامة بين الظاهرتين س ، ص . وهى حالة مخالفة للحالة السابقة تماماً ، حيث نجد أن الظاهرتين س ، ص تسيران فى اتجاهين متضادين . وهذا يعنى أن أكبر قيمة للظاهرة س تناظرها أصغر قيمة للظاهرة ص . والقيمة الثانية فى الكبر للظاهرة س تناظرها القيمة الثانية فى الصغر للظاهرة ص . وهكذا نجد أنه بالاستمرار فى التدريج نصل فى النهاية إلى أن تكون أصغر قيمة للظاهرة س تناظرها أكبر قيمة للظاهرة ص

ثالثاً : معامل الارتباط يقع بين + ١ ، - ١ :

إذا كانت العلاقة بين المتغيرين س ، ص لا هى بالطردية التامة ولا هى بالعكسية التامة . وتحدد القيمة العددية لمعامل الارتباط درجة قوة أو ضعف هذه العلاقة . فكلما اقترب المعامل من الواحد الصحيح كلما دل ذلك على قوة العلاقة بين المتغيرين . وبالعكس من ذلك إذ كلما اقترب هذا المعامل من الصفر كلما دل ذلك على ضعف العلاقة بين المتغيرين .

أما إشارة معامل الارتباط فهى تعين اتجاه العلاقة بين المتغيرين . فتكون الاشارة موجبة إذا كانت العلاقة طردية ، وتكون سالبة إذا كانت العلاقة عكسية .

وعلى ضوء ذلك فإن معامل الارتباط يكون أكبر من الصفر ولكن أصغر من الواحد الصحيح حين تكون العلاقة بين المتغيرين س ، ص طردية ناقصة . ويكون معامل الارتباط أكبر من - ١ وأصغر من الصفر حين تكون العلاقة بين المتغيرين س ، ص عكسية ناقصة .

هذا ووقوع معامل الارتباط بين + ١ ، - ١ تمثل الحالات الأكثر شيوعاً فى كثير من الظواهر الاقتصادية أو الإدارية أو الاجتماعية مثل العلاقة بين الكمية المطلوبة أو المعروضة من سلعة معينة وسعرها ، والعلاقة بين الإعلان وكمية المبيعات من سلعة معينة ، والعلاقة بين الحالة التعليمية لمجموعة معينة من الأشخاص ودخلهم ، والعلاقة بين الطول والوزن لمجموعة معينة من الأشخاص فى سن معين . . الخ .

هذا ويجب ملاحظة أن كون معامل الارتباط مختلفاً عن الصفر ، لا يعنى بالضرورة وجود علاقة سببية مباشرة بين المتغيرين . فقد يكون ذلك ناتجاً عن تأثير عامل مشترك أو عدة عوامل مشتركة أخرى أدت إلى وجود هذا الارتباط بينهما فمثلا وجود ارتباط بين دخل الزوج ودخل الزوجة لا يعنى أن يكون أيهما سبب للآخر فقد يكون هناك عاملاً مشتركاً هو السبب فى وجود هذا الارتباط وليكن الحالة التعليمية لكل

منهما . وبالمثل وجود ارتباط بين محصول القصب ومحصول الأرز خلال عدد معين من السنوات لا يعني أن يكون أحدهما سبب للآخر أو منشئ له ، فقد يرجع هذا الارتباط بينهما إلى وجود عامل مشترك يتأثران به ، وليكن هذا العامل هو وفرة مياه الري .

وبالمثل وجود ارتباط بين درجات الرياضة البحته ودرجات التاريخ في امتحان ما لايعني أن يكون أحداهما سبب للآخر لما بينهما من تباعد كبير ، فقد يرجع الارتباط بينهما إلى أن كليهما يتأثران بعامل مشترك وليكن قراءة الأسئلة واستيعابها .

رابعاً : معامل الارتباط يساوي الصفر :

إذا كانت العلاقة بين المتغيرين س ، ص منعدمة . مثال ذلك العلاقة بين طول الشخص ودخله ، والعلاقة بين لون الشعر ودرجة الذكاء .

طرق حساب معامل بيرسون للارتباط

أولاً : البيانات غير المبوبة

أولاً :الطريقة المطولة باستخدام الوسط الحسابي لكل من الظاهرتين :

حساب معامل الارتباط من الصيغة :

$$ر = \frac{\text{مجـ}(س - \bar{س})(ص - \bar{ص})}{\sqrt{\text{مجـ}(س - \bar{س})^2}\sqrt{\text{مجـ}(ص - \bar{ص})^2}} \qquad (1)$$

يعرف بالطريقة المطولة . وتتلخص طريقة الحساب في اتباع الخطوات الآتية وذلك بعد تدوين قيم المتغير س في العمود الأول وقيم المتغير ص في العمود الثاني من الجدول :

١ – نحسب الوسط الحسابي لكل من المتغيرين س ، ص حيث :

$$\bar{س} = \frac{\text{مجـ}س}{ن} \quad ، \quad \bar{ص} = \frac{\text{مجـ}ص}{ن}$$

٢ – نوجد انحراف كل قيمة من قيم المتغير س عن وسطها الحسابي $\bar{س}$ ، وندون الانحرافات الناتجة في العمود الثالث وبالمثل نوجد انحراف كل قيمة من قيم المتغير ص عن وسطها الحسابي $\bar{ص}$ ، وندون الانحرافات الناتجة في العمود

الرابع . وطبقاً لخصائص الوسط الحسابي فإن مجموع انحرافات القيم عنه لابد وأن يساوى الصفر ويترتب على ذلك أن مجـ (س - $\overline{س}$) = صفراً ، مجـ (ص - $\overline{ص}$) = صفراً .

٣ - لايجاد مقام كسر معامل الارتباط فإننا نربع انحرافات كل قيمة من المتغيرين س ، ص والواردين فى العمود الثالث والعمود الرابع . ونـدون مربعات الانحرافات الخاصة بالمتغير س فى العمود الخامس ، ومربعات الانحرافات الخاصة بالمتغير ص فى العمود السادس ، ثم بجمع كل عمود من هذين العمودين نحصل على مجـ (س - $\overline{س}$)^{٢} ، مجـ (ص - $\overline{ص}$)^{٢}

٤ - لإيجاد بسط معامل الارتباط . نضرب كل انحراف من انحرافات قيم المتغير س والواردة فى العمود الثالث فى الانحرافات المناظرة لها للمتغير ص والواردة فى العمود الرابع ، ثم ندون حواصل الضرب الناتجة فى العمود السابع والأخير من الجدول . ثم يجمع هذا العمود نحصل على مجـ (س - $\overline{س}$) (ص - $\overline{ص}$) .

ولإيجاد معامل الارتباط نعوض بالقيم السابقة فى المعادلة (١) .

مثال (١) :

الجدول الآتى يعطى درجات عشرة تلاميذ فى مادة الرياضة البحتة (س) والاقتصاد (ص) :

س	١٨	١٤	١٢	٨	١٢	١١	١٦	١٥	١٦	٨
ص	١٩	١٥	١١	٧	١١	٩	١٣	١٦	١٧	٧

والمطلوب حساب معامل بيرسون للارتباط الخطى البسيط

الحـل : ننشئ الجدول الآتي :

س	ص	$(س - \bar{س})$ (س - ١٢)	$(ص - \bar{ص})$ (ص - ١٢,٥)	$(س - \bar{س})^2$ $(س - ١٢)^2$	$(ص - \bar{ص})^2$ $(ص - ١٢,٥)^2$	$(س - \bar{س})(ص - \bar{ص})$ (س - ١٢)(ص - ١٢,٥)
١٨	١٩	٥	٦,٥	٢٥	٤٢,٢٥	٣٢,٥
١٤	١٥	١	٢,٥	١	٦,٢٥	٢,٥
١٢	١١	١-	١,٥-	١	٢,٢٥	١,٥
٨	٧	٥-	٥,٥-	٢٥	٣٠,٢٥	٢٧,٥
١٢	١١	١-	١,٥-	١	٢,٢٥	١,٥
١١	٩	٢-	٣,٥-	٤	١٢,٢٥	٧,٠
١٦	١٣	٣	٠,٥	٩	٠,٢٥	١,٥
١٥	١٦	٢	٣,٥	٤	١٢,٢٥	٧,٠
١٦	١٧	٣	٤,٥	٩	٢٠,٢٥	١٣,٥
٨	٧	٥-	٥,٥-	٢٥	٣٠,٢٥	٢٧,٥
١٣٠	١٢٥	١٤ ١٤- صفر	١٧,٥ ١٧,٥- صفر	١٠٤	١٥٨,٥٠	١٢٢,٠

نحسب أولاً الوسط الحسابى لكل من المتغيرين س ، ص حيث

$$\bar{س} = \frac{مجـ س}{ن} = \frac{١٣٠}{١٠} = ١٣$$

$$\bar{ص} = \frac{مجـ ص}{ن} = \frac{١٢٥}{١٠} = ١٢.٥$$

من الجدول السابق نجـــد أن :

$$مجـ (س - \bar{س})(ص - \bar{ص}) = مجـ (س - ١٣)(ص - ١٢.٥) = ١٢٢$$

$$، مجـ (س - \bar{س})^٢ = مجـ (س - ١٣)^٢ = ١٠٤$$

$$، مجـ (ص - \bar{ص})^٢ = مجـ (ص - ١٢.٥)^٢ = ١٥٨.٥$$

وبالتعويض بهذه القيم فى معادلة " معامل بيرسون للارتباط الخطى البسيط " (١) نجد أن :

$$ر = \frac{١٢٢}{\sqrt{١٠٤}\sqrt{١٥٨.٥}} = \frac{١٢٢}{\sqrt{(١٠٤)(١٥٨.٥)}}$$

$$= \frac{١٢٢}{\sqrt{١٦٤٨٤}} = \frac{١٢٢}{١٢٨.٣٩} = ٠.٩٥$$

وهذا يعنى أن العلاقة بين المتغيرين س ، ص طردية وقوية جداً ، حيث أن قيمة معامل الارتباط تقترب من الواحد الصحيح .

ثانياً :الطريقة المباشرة باستخدام القيم الأصلية لكل من الظاهرتين :

يفضل استخدام القيم الأصلية لكل من المتغيرين س ، ص إذا ما كانت هذه القيم صغيرة نسبياً وكانت عملية حساب الارتباط لا تتطلب عمليات حسابية معقدة ،وفى هذه الحالة يستلزم الأمر تعديل صيغة معامل الارتباط السابقة بما يتناسب مع استخدام هذه الطريقة وذلك باجراء بعض العمليات الجبرية .

<u>وتكون صيغة معامل الارتباط البسيط باستخدام القيم الأصلية للمتغيرين</u>

س ، ص على الوجه الآتى :

$$\text{ر} = \frac{\text{مجـ س ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}}{\sqrt{\text{مجـ س}^{2} - \text{ن}\,\bar{\text{س}}^{2}}\;\sqrt{\text{مجـ ص}^{2} - \text{ن}\,\bar{\text{ص}}^{2}}} \qquad (٢)$$

مثال (٢) :

سوف نقوم بحل المثال (١) بطريقة القيم الأصلية للمتغيرين :

ننشئ الجدول الآتى :

س	ص	س^2	ص^2	س ص
١٨	١٩	٣٢٤	٣٦١	٣٤٢
١٤	١٥	١٩٦	٢٢٥	٢١٠
١٢	١١	١٤٤	١٢١	١٣٢
٨	٧	٦٤	٤٩	٥٦
١٢	١١	١٤٤	١٢١	١٣٢
١١	٩	١٢١	٨١	٩٩
١٦	١٣	٢٥٦	١٦٩	٢٠٨
١٥	١٦	٢٢٥	٢٥٦	٢٤٠
١٦	١٧	٢٥٦	٢٨٩	٢٧٢
٨	٧	٦٤	٤٩	٥٦
١٣٠	١٢٥	١٧٩٤	١٧٢١	١٧٤٧

من الجدول السابق نجد أن

$$\bar{\text{س}} = \frac{\text{مجـ س}}{\text{ن}} = \frac{\text{١٣٠}}{\text{١٠}} = \text{١٣}$$

$$\bar{\text{ص}} = \frac{\text{مجـ ص}}{\text{ن}} = \frac{\text{١٢٥}}{\text{١٠}} = \text{١٢.٥}$$

وبالتعويض فى المعادلة (٢) نجد أن :

$$r = \frac{\text{١٧٤٧} - \text{١٠}(\text{١٣})(\text{١٢.٥})}{\sqrt{\text{١٧٩٤} - \text{١٠}(\text{١٣})^{\text{٢}}}\ \sqrt{\text{١٧٢١} - (\text{١٢.٥})^{\text{٢}}}}$$

$$= \frac{\text{١٧٤٧} - \text{١٦٢٥}}{\sqrt{\text{١٧٩٤} - \text{١٦٩٠}}\ \sqrt{\text{١٧٢١} - \text{١٥٦٢.٥}}}$$

$$= \frac{\text{١٢٢}}{\sqrt{\text{١٠٤}}\ \sqrt{\text{١٥٨.٥}}} = \frac{\text{١٢٢}}{\sqrt{(\text{١٠٤})(\text{١٥٨.٥})}}$$

$$= \frac{\text{١٢٢}}{\sqrt{\text{١٦٤٨.٤}}} = \frac{\text{١٢٢}}{\text{١٢٨.٣٩}} = \underline{\underline{\text{٠.٩٥}}}$$

وهى نفس النتيجة السابقة .

ثالثاً :طريقة الانحرافات باستخدام وسط فرضى لكل من المتغيرين :

كثيراً ما يتضمن الوسط الحسابى لكل من أو لإحدى المتغيرين عدداً كسرياً . ونظراً لأن حساب الارتباط باستخدام الطريقة المطولة يستلزم حساب انحراف كل قيمة من قيم المتغير عن الوسط الحسابى الخاص به ، ثم تربيع هذه الانحرافات ، لذلك فإن الأمر يتطلب إجراء عمليات حسابية مطولة ومعقدة للغاية فى مثل هذه الحالات .

هذا ويمكن تبسيط العمليات الحسابية عن طريق أخذ وسط فرضى لكل من المتغيرين س ، ص وتعرف هذه الطريقة بطريقة الانحرافات ، حيث :

$$r = \frac{\text{مجـ } \text{ح}_{\text{س}}\text{ح}_{\text{ص}} - \text{ن}\,\bar{\text{ح}}_{\text{س}}\bar{\text{ح}}_{\text{ص}}}{\sqrt{\text{مجـ } \text{ح}^{\text{٢}}_{\text{س}} - \text{ن}\,\bar{\text{ح}}^{\text{٢}}_{\text{س}}}\ \sqrt{\text{مجـ } \text{ح}^{\text{٢}}_{\text{ص}} - \text{ن}\,\bar{\text{ح}}^{\text{٢}}_{\text{ص}}}} \qquad (\text{٣})$$

مثال (٣) : سوف نقوم بحل المثال (١) باستخدام وسطاً فرضياً للمتغير س وقدره ١٤ ، ووسطاً فرضياً للمتغير ص وقدره ١٣ :

س	ص	(س - و) = ح س (س - ١٤)	(ص - وَ) = ح ص (ص - ١٢)	ح^{2} س (س - ١٤)^{2}	ح^{2} ص (ص - ١٢)^{2}	ح س ح ص (س - ١٤) (ص - ١٢)
١٨	١٩	٤	٦	١٦	٣٦	٢٤
١٤	١٥	صفر	٢	صفر	٤	صفر
١٢	١١	٢-	٢-	٤	٤	٤-
٨	٧	٦-	٦-	٣٦	٣٦	٣٦
١٢	١١	٢-	٢-	٤	٤	٤
١١	٩	٣-	٤-	٩	١٦	١٢
١٦	١٣	٢	صفر	٤	صفر	صفر
١٥	١٦	١	٣	١	٩	٣
١٦	١٧	٢	٤	٤	١٦	٨
٨	٧	٦-	٦-	٣٦	٣٦	٣٦
١٣٠	١٢٥	١٩- ٩ ١٠-	٢٠- ١٥ ٥-	١١٤	١٦١	١٢٧

من الجدول السابق نجد أنه بالنسبة لبسط المعادلة (٣) نجد أن :

$$\text{مجـ } ح_{س} \, ح_{ص} = ١٢٧$$

$$\bar{ح}_{س} = \frac{\text{مجـ } ح_{س}}{ن} = \frac{-١٠}{١٠} = -١$$

$$\bar{ح}_{ص} = \frac{\text{مجـ } ح_{ص}}{ن} = \frac{-٥}{١٠} = -٠.٥$$

وبالنسبة لمقام نفس المعادلة فإننا نجد أن :

$$ح^{٢}_{س} = ١١٤ \qquad\qquad ح^{٢}_{ص} = ١٦١$$

وعلى ذلك يكون قيمة معامل الارتباط الخطى البسيط باستخدام وسط فرضى لكل

من المتغيرين س ، ص طبقاً للمعادلة (٣) كالآتى :

$$ر = \frac{١٢٧ - ١٠(-١)(-٠.٥)}{\sqrt{١١٤ - ١٠(-١)^{٢}}\sqrt{١٦١ - ١٠(-٠.٥)^{٢}}}$$

$$= \frac{١٢٧ - ٥}{\sqrt{١١٤ - ١٠}\sqrt{١٦١ - ٢.٥}} = \frac{١٢٢}{\sqrt{١٠٤}\sqrt{١٥٨.٥}}$$

$$= \frac{١٢٢}{\sqrt{(١٠٤)(١٥٨.٥)}} = \frac{١٢٢}{\sqrt{١٦٤٨.٤}}$$

$$= \frac{١٢٢}{١٢٨.٣٩} = ٠.٩٥$$

وهى نفس النتيجة السابقة

وهذه النتيجة يمكن ملاحظاتها من مقارنة المعادلة (٢) بالمعادلة (٣)

حيث قد استبدل المتغيران س ص بانحرافاتهما عن الوسط الفرضى الخاص

بكل منها . أما المتوسطان $\bar{س}$ ، $\bar{ص}$ فقد استبدلا بمتوسطات الانحرافات $\bar{ح}_{س}$ ، $\bar{ح}_{ص}$ وهذا يعنى أنه سواء استخدمنا القيم الأصلية للمتغيرين أو استخدمنا الانحرافات عن وسط فرضى مناسب لكل منهما فاننا نحصل على نفس قيمة معامل الارتباط

مثال (٤) :

الجدول الآتى يوضح الطول بالبوصة (س) والوزن بالرطل (ص) لعشرة طلاب من كلية معينة .

س	٦٤	٧١	٧٠	٦٨	٧٠	٦٨	٧٣	٦٩	٧٠	٦٧
ص	١٥٦	١٧٨	١٧٥	١٦٧	١٦٤	١٦٠	١٩٨	١٦٩	١٨٣	١٨٠

والمطلوب حساب معامل بيرسون للارتباط باستخدام :

أولاً : الوسط الحسابى لكل من المتغيرين .

ثانياً : الوسط الفرضى لكل من المتغيرين .

ثالثاً : القيم الأصلية للمتغيرين .

الحـــل :

أولاً : باستخدام الوسط الحسابى لكل من المتغيرين :

من الجدول الآتى يتضح أن

$$\bar{س} = \frac{مجـ\ س}{ن} = \frac{٦٩٠}{١٠} = ٦٩$$

$$\bar{ص} = \frac{مجـ\ ص}{ن} = \frac{١٧٣}{١} = ١٧٣$$

الحل نتشئ الجدول الآتي :

س	ص	(س - $\bar{س}$) (س - ٦٩)	(ص - $\bar{ص}$) (ص - ١٧٣)	$(س - \bar{س})^2$ $(س - ٦٩)^2$	$(ص - \bar{ص})^2$ $(ص - ١٧٣)^2$	(س - $\bar{س}$) (ص - $\bar{ص}$) (س - ٦٩) (ص - ١٧٣)
٦٤	١٥٦	٥-	١٧-	٢٥	٢٨٩	٨٥
٧١	١٧٨	٢	٥	٤	٢٥	١٠
٧٠	١٧٥	١	٢	١	٤	٢
٦٨	١٦٧	١-	٦-	١	٣٦	٦
٧٠	١٦٤	١	٩-	١	٨١	٩-
٦٨	١٦٠	١-	١٣-	١	١٦٩	١٣
٧٣	١٩٨	٤	٢٥	١٦	٦٢٥	١٠٠
٦٩	١٦٩	٠	٤-	٠	١٦	٠
٧٠	١٨٣	١	١٠	١	١٠٠	١٠
٦٧	١٨٠	٢-	٧	٤	٤٩	١٤-
٦٩٠	١٧٣٠	٩ ٩- صفر	٤٩ ٤٩- صفر	٥٤	١٣٩٤	٢٢٦ ٢٣- ٢٠٣

من الجدول السابق نجد أن .

$$\text{مجـ}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}}) = \text{مجـ}(\text{س}-٦٩)(\text{ص}-١٧٣) = ٢٠٣$$

$$\text{مجـ}(\text{س}-\bar{\text{س}})^{٢} = \text{مجـ}(\text{س}-٦٩)^{٢} = ٥٤$$

$$\text{مجـ}(\text{ص}-\bar{\text{ص}})^{٢} = \text{مجـ}(\text{ص}-١٧٣)^{٢} = ١٣٩٤$$

وبالتعويض بهذه القيم فى المعادلة ( ) الخاصة بحساب معامل بيرسون للارتباط الخطي البسيط وهى :

$$\text{ر} = \frac{\text{مجـ}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}})}{\sqrt{\text{مجـ}(\text{س}-\bar{\text{س}})^{٢}}\sqrt{\text{مجـ}(\text{ص}-\bar{\text{ص}})^{٢}}} \qquad (1)$$

$$= \frac{٢٠٣}{\sqrt{٥٤}\sqrt{١٣٩٤}} = \frac{٢٠٣}{\sqrt{٥٤\times١٣٩٤}}$$

$$= \frac{٢٠٣}{\sqrt{٧٥٢٧٦}} = \frac{٢٠٣}{٢٧٤.٣٦} = ٠.٧٤$$

وهذا يعنى أن العلاقة بين المتغيرين س ، ص طردية وقوية .

ثانياً :الحل باستخدام طريقة الانحرافات لكل من المتغيرين س، ص

خذ وسطاً فرضياً للمتغير س = ٧٠

للمتغير ص = ١٧٥

الحل : ننشئ الجدول الآتي :

س	ص	ح س (س − ٧٠)	ح ص (ص − ١٧٥)	ح²س (س − ٧٠)²	ح²ص (ص − ١٧٥)²	ح س ح ص (س − ٧٠) (ص − ١٧٥)
٦٤	١٥٦	−٦	−١٩	٣٦	٣٦١	١١٤
٧١	١٧٨	١	٣	١	٩	٣
٧	١٧٥	٠	٠	٠	٠	٠
٦٨	١٦٧	−٢	−٨	٤	٦٤	١٦
٧٠	١٦٤	٠	−١١	٠	١٢١	٠
٦٨	١٦٠	−٢	−١٥	٤	٢٢٥	٣٠
٧٣	١٩٨	٣	٢٣	٩	٥٢٩	٦٩
٦٩	١٦٩	−١	−٦	١	٣٦	٦
٧٠	١٨٣	٠	٨	٠	٦٤	٠
٦٧	١٨٠	−٣	٥	٩	٢٥	−١٥
٦٩٠	١٧٣٠	−١٤ ٤ ——— −١٠	−٥٩ ٣٩ ——— −٢٠	٦٤	١٤٢٤	٢٢٨ −١٥ ——— ٢٢٣

وحيث أن معادلة حساب معامل الارتباط الخطى البسيط باستخدام وسط فرضى (٢) تأخذ الصيغة الآتية :

$$\text{ر} = \frac{\text{مجـ } \text{ح}_{\text{س}} \text{ح}_{\text{ص}} - \text{ن} \, \bar{\text{ح}}_{\text{س}} \bar{\text{ح}}_{\text{ص}}}{\sqrt{\text{مجـ } \text{ح}^2_{\text{س}} - \text{ن} \, \bar{\text{ح}}^2_{\text{س}}} \sqrt{\text{مجـ } \text{ح}^2_{\text{ص}} - \text{ن} \, \bar{\text{ح}}^2_{\text{ص}}}} \qquad (٢)$$

فإنه من الجدول السابق ، فإن بسط المعادلة السابقة يتطلب استخدام القيم الآتية :

$$\text{مجـ } \text{ح}_{\text{س}} \text{ح}_{\text{ص}} = \text{مجـ } (\text{س} - ٧٠)(\text{ص} - ١٧٥) = ٢٢٣$$

$$\therefore \bar{\text{ح}}_{\text{س}} = \frac{\text{مجـ } \text{ح}_{\text{س}}}{\text{ن}} = \frac{-١٠}{١٠} = -١$$

$$\therefore \bar{\text{ح}}_{\text{ص}} = \frac{\text{مجـ } \text{ح}_{\text{ص}}}{\text{ن}} = \frac{-٢٠}{١٠} = -٢$$

وبالنسبة لمقام نفس المعادلة فإننا نجد أن :

$$\text{ح}^2_{\text{س}} = ٦٤ \quad ، \quad \text{ح}^2_{\text{ص}} = ١٤٣٤$$

وعلى ذلك يحسب معامل الارتباط الخطى البسيط كالآتى :

$$\text{ر} = \frac{٢٢٣ - ١٠(-١)(-٢)}{\sqrt{٦٤ - ١٠(-١)^2} \sqrt{١٤٣٤ - ١٠(-٢)^2}}$$

$$= \frac{٢٢٣ - ٢٠}{\sqrt{٦٤ - ١٠} \sqrt{١٤٣٤ - ٤٠}} = \frac{٢٠٣}{\sqrt{٥٤} \sqrt{١٣٩٤}}$$

$$= \frac{٢٠٣}{\sqrt{٧٥٢٧٦}} = \frac{٢٠٣}{٢٧٤.٣٦} = ٠.٧٤$$

وهى نفس النتيجة السابقة

ثالثاً : باستخدام القيم المباشرة للمتغيرين س ، ص :

س	ص	س٢	ص٢	س ص
٦٤	١٥٦	٤٠٩٦	٢٤٣٣٦	٩٩٨٤
٧١	١٧٨	٥٠٤١	٣١٦٨٤	١٢٦٣٨
٧٠	١٧٥	٤٩٠٠	٣٠٦٢٥	١٢٢٥٠
٦٨	١٦٧	٤٦٢٤	٢٧٨٨٩	١١٣٥٦
٧٠	١٦٤	٤٩٠٠	٢٦٨٩٦	١١٤٨٠
٦٨	١٦٠	٤٦٢٤	٢٥٦٠٠	١٠٨٨٠
٧٣	١٩٨	٥٣٢٩	٣٩٢٠٤	١٤٤٥٤
٦٩	١٦٩	٤٧٦١	٢٨٥٦١	١١٦٦١
٧٠	١٨٣	٤٩٠٠	٣٣٤٨٩	١٢٨١٠
٦٧	١٨٠	٤٤٨٩	٣٢٤٠٠	١٢٠٦٠
٦٩٠	١٧٣٠	٤٧٦٦٤	٣٠٠٦٨٤	١١٩٥٧٣

وحيث أن معادلة حساب معامل الارتباط الخطى البسيط باستخدام القيم الأصلية للمتغيرين س ، ص تأخذ الصيغة (٣) الآتية :

$$\text{ر} = \frac{\text{مجـ س ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}}{\sqrt{\text{مجـ س}^2 - \text{ن}\,\bar{\text{س}}^2}\;\sqrt{\text{مجـ ص}^2 - \text{ن}\,\bar{\text{ص}}^2}} \qquad (٣)$$

وبالاستعانة بالعمليات الحسابية بالجدول أعلاه نجد أنه بالنسبة لبسط معامل الارتباط نجد أن مجـ س ص = ١١٩٥٧٣

$$\bar{\text{س}} = \frac{\text{مجـ س}}{\text{ن}} = \frac{٦٩٠}{١٠} = ٦٩$$

$$\bar{\text{ص}} = \frac{\text{مجـ ص}}{\text{ن}} = \frac{١٧٣٠}{١٠} = ١٧٣$$

وبالنسبة لمقام معامل الارتباط نجد أن :

$$\text{مجـ س}^{2} = ٤٧٦٦٤ \quad ، \quad \text{مجـ ص}^{2} = ٢٠٠٦٨٤$$

وعلى ذلك يحسب معامل الارتباط الخطى البسيط كالآتى : -

$$ر = \frac{١١٩٥٧٣ - ١٠ (٦٩)(١٧٣)}{\sqrt{٤٧٦٦٤ - ١٠ (٦٩)^{2}} \sqrt{٢٠٠٦٨٤ - ١٠ (١٧٣)^{2}}}$$

$$= \frac{١١٩٥٧٣ - ١١٩٣٧٠}{\sqrt{٤٧٦٦٤ - ٤٧٦١٠} \sqrt{٢٠٠٦٨٤ - ١٩٩٢٩٠}} = \frac{٢٠٣}{\sqrt{٥٤} \sqrt{١٣٩٤}}$$

$$= \frac{٢٠٣}{\sqrt{٧٥٢٧٦}} = \frac{٢٠٣}{٢٧٤.٣٦} = ٠.٧٤$$

ملاحظات على طرق حساب معامل الارتباط :

١ - إذا كان الوسط الحسابى لكل من المتغيرين س ، ص عددين صحيحين وكانت القيم العددية لهما كبيرة ، فإنه يفضل فى هذه الحالة حساب معامل الارتباط باستخدام هذين الوسطين الحسابين إذ أن ذلك يؤدى إلى تبسيط قيم المتغيرين وبالتالى تبسيط عملية حساب معامل الارتباط ، فضلا عن أن هذه الطريقة تمكننا من التأكد من صحة هذه القيم المبسطة (الانحرافات عن الوسط الحسابى) إذ أن مجموعها لابد وأن يساوى الصفر طبقاً لخصائص الوسط الحسابى .

٢ - إذا كانت قيم المتغيرين س ، ص صغيرة نسبياً بحيث يكون من السهولة بمكان إيجاد المربعات وحواصل الضرب لهذه القيم دون أن يتطلب الأمر عمليات حسابية مطولة ، فإنه يحسن استخدام قيم المتغيرين مباشرة لما فيها من اختصار فى العمل لعدم تطلبها استخراج الانحرافات سواء عن طريق الوسط الحسابى أو الفرضى

٣ ـ إذا احتوى الوسط الحسابى لإحدى أو لكلا المتغيرين على عدد كسرى فإن استخدام الطريقة المطولة بإيجاد انحراف كل قيمة من قيم المتغير عن وسطها الحسابى ، ثم تربيع هذه الإنحرافات ، ثم إيجاد حواصل الضرب ، فيه تعقيد كبير فى العمليات الحسابية ويفضل فى هذه الحالة استخدام وسط فرضى مناسب إذ أن ذلك يؤدى إلى تبسيط العمليات الحسابية .

مثال (٥) :

من البيانات الآتية احسب معامل الارتباط الخطى البسيط للمتغيرين س ، ص .

س	١٣	١٤	١٥	١٦	١٧	١٩	٢٢	٢٥
ص	١٩	١٨	١٧	١٦	١٥	١٥	١٤	١٣

الحل :

$$\bar{س} = \frac{مجـ س}{ن} = \frac{١٤١}{٨} = ١٧,٦٢٥$$

$$\bar{ص} = \frac{مجـ ص}{ن} = \frac{١٢٧}{٨} = ١٥,٨٧٥$$

لذلك فإن استخدام الطريقة المطولة ، عن طريق حساب انحرافات قيم كل من المتغيرين س ، ص عن وسطها الحسابى الكسرى ، ثم تربيع هذه الانحرافات ، ثم ايجاد حواصل الضرب يتطلب عمليات حسابية مطولة ، لذلك لا يفضل استخدام هذه الطريقة فى حل التمرين

ومن المفضل فى هذه الحالة استخدام قيم المتغيرين مباشرة لما فيها من اختصار فى العمل لعدم تطلبها استخراج الانحرافات سواء عن طريق الوسط الحسابى أو الفرضى للمتغيرين س ص .

س	ص	س٢	ص٢	س ص
١٣	١٩	١٦٩	٣٦١	٢٤٧
١٤	١٨	١٩٦	٣٢٤	٢٥٢
١٥	١٧	٢٢٥	٢٨٩	٢٥٥
١٦	١٦	٢٥٦	٢٥٦	٢٥٦
١٧	١٥	٢٨٩	٢٢٥	٢٥٥
١٩	١٥	٣٦١	٢٢٥	٢٨٥
٢٢	١٤	٤٨٤	١٩٦	٣٠٨
٢٥	١٣	٦٢٥	١٦٩	٣٢٥
١٤١	١٢٧	٢٦٠٥	٢٠٤٥	٢١٨٣

وقد سبق أن ذكرنا صيغة معامل الارتباط الخطى البسيط باستخدام الصيغة المباشرة (أى القيم الأصلية للمتغيرين) هى :

$$\text{ر} = \frac{\text{مجـ س ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}}{\sqrt{\text{مجـ س}^2 - \text{ن}\,\bar{\text{س}}^2}\ \sqrt{\text{مجـ ص}^2 - \text{ن}\,\bar{\text{ص}}^2}} \qquad (٣)$$

$$= \frac{٢١٨٣ - ٨(١٧.٦٢٥)(١٥.٨٧٥)}{\sqrt{٢٦٠٥ - ٨(١٧.٦٢٥)^2}\ \sqrt{٢٠٤٥ - ٨(١٥.٨٧٥)^2}}$$

$$= \frac{٢١٨٣ - ٢٢٣٨.٣٧٥}{\sqrt{٢٦٠٥ - ٢٤٨٥.١٢٥}\ \sqrt{٢٠٤٥ - ٢٠١٦.١٢٥}} = \frac{-٥٥.٣٧٥}{\sqrt{١١٩.٨٧٥ \times ٢٨.٨٧٥}}$$

$$= \frac{-٥٥.٣٧٥}{\sqrt{٣٤٦١.٣٩٠٦}} = \frac{-٥٥.٣٧٥}{٥٨.٨٣٣٦} = -٠.٩٤$$

من الواضح أن معامل الارتباط الخطى البسيط ذو كسر كبير يقرب من الواحد الصحيح مما يوضح وجود علاقة قوية جداً وعكسية (سالبة) بين المتغيرين س ص

ثانياً : البيانات المبوبة

قد تكون مجموعة أزواج القيم الخاصة بالظاهرتين س ، ص موضوع الدراسة كبيرة جداً ، بحيث تكون من الصعوبة بمكان إجراء العمليات الحسابية الخاصة بالتحليل الاحصائى عليها ، لذلك فإن الأمر يستلزم فى مثل هذه الحالات عرض بيانات هاتين الظاهرتين فى صورة مختصرة تسهل معها عملية التحليل . ويتم ذلك بتبويب هذه البيانات فى صورة جدول تكرارى مزدوج ، بحيث يخصص فيه الصف الأول لفئات إحدى المتغيرين والعمود الأول لفئات المتغير الأخر . وتمثل القيم الواردة بخلايا هذا الجدول التكرارات المشتركة المناظرة لإحدى فئات المتغير س ولإحدى فئات المتغير ص .

فمثلاً الجدول الآتى يعطى أوزان ٥٠ شخصاً بالكيلوجرامات (س) وأطوالهم مقيسة بالسنتيمترات (ص) :

ص \ س	٦٢-	٦٦-	٧٠-	٧٤-	٧٨-٨٢	ك ص
١٦٠-	٢	٢				٤
١٦٥-	٣	٦	٣			١٢
١٧٠-		٤	٨	٢		١٤
١٧٥-			٦	٤		١٠
١٨٠-			١	٥	١	٧
١٨٥-١٩٠				٢	١	٣
ك س	٥	١٢	١٨	١٣	٢	٥٠

من هذا الجدول يتضح أن الصف الأول منه يوضح فئات المتغير س . أما مجاميع الأعمدة الرأسية المناظرة لهذه الفئات والواردة فى الصف الأخير المعنون بـ $ك_{س}$ والتى تسمى بالمجاميع الهامشية marginal totals فتمثل التكرارات الخاصة بها . وعلى ذلك يمكننا تصوير جدول التوزيع التكرارى البسيط للمتغير س أو ما يعرف بجدول التوزيع الهامشى للمتغير س (الأوزان بالكيلو جرامات) كالآتى : -

فئـــــات س	٦٢-	٦٦-	٧٠-	٧٤-	٧٨ - ٨٢	المجموع
التكرارات : $ك_{س}$	٥	١٢	١٨	١٢	٣	٥٠

وبالمثل يتضح أن العمود الأول من جدول التوزيع التكرارى المزدوج المشار إليه يوضح فئات المتغير ص ، أما مجاميع الصفوف الأفقية المناظرة لهذه الفئات الواردة فى العمود الأخير المعنون بـ $ك_{ص}$ فتمثل التكرارات الخاصة بها ، وعلى ذلك يمكننا تصوير جدول التوزيع التكرارى البسيط للمتغير ص أو ما يعرف بجدول التوزيع الهامشى للمتغير ص (الأطوال بالسنتيمترات) كالآتى :

فئـــــات ص	١٦٠-	١٦٥-	١٧٠-	١٧٥-	١٨٠-	١٨٥ - ١٩٠	المجموع
التكرارات : $ك_{ص}$	٤	١٢	١٤	١٠	٧	٣	٥٠

أما بالنسبة لخلايا الجدول فكما أوضحنا أن كل منها تمثل التكرارات المشتركة بين المتغيرين س ، ص فمثلاً الخلية الأولى من الصف الأول تعنى أن هناك شخصين تتراوح أوزانهم بالكيلوجرامات بين ٦٢ - ٦٦ وأطوالهم بالسنتيمترات بين ١٦٠ - ١٦٥ ونظراً لأننا نعبر عن كل من فئـــات س ، ص بمراكـــزها ، فإننا نقول أن هـــناك تكرارين لكل منها س تأخذ القيمة ٦٤ ولكل منها ص تأخذ ١٦٢.٥ .

هذا ويمكن اكتشاف طبيعة العلاقة بين الظاهرتين س ، ص والمبوبة فى صورة جدول توزيع مزدوج ، من فحص التكرارات الواقعة داخل خلايا هذا الجدول . فإذا كانت التكرارات متركزة حول قطر الجدول الممتد من أدنى اليسار إلى أعلى اليمين ، فإن هذا يعنى أن قيم س الكبيرة تميل إلى أن تصاحبها قيم ص الكبيرة ، وأن قيم س الصغيرة تميل إلى أن تصاحبها قيم ص الصغيرة ، وهذا يعنى وجود علاقة طردية (موجبة) بين المتغيرين . وتتوقف درجة قوة أو ضعف هذه العلاقة على مدى تركز هذه التكرارات حول هذا القطر .

وبالعكس من ذلك إذا كانت التكرارات متركزة حول القطر الممتد من أعلى اليسار إلى أدنى اليمين ، وهذه حالة مخالفة للسابقة ، فإن هذا يعنى وجود علاقة عكسية بين المتغيرين . وتتوقف درجة قوة أو ضعف هذه العلاقة على مدى تكز هذه التكرارت حول هذا القطر . .

هذا وإذا كانت التكرارات مشتتة فى جميع خلايا الجدول ، كلما دل ذلك على ضعف أو عدم وجود علاقة بين المتغيرين موضوع الدراسة .

وعلى ضوء ما تقدم فإن تجمــع التكرارات وشكل هذا التجمع بجدول التوزيع التكرارى المزدوج ، يكشـف لنا بطـريقة تقريبية عن طبيعـة العلاقة التى تربط بين المتغــيرين .

هذا وإذا ما تبين أن العلاقة بين المتغيرين خطية ، فتكون الخطوة التالية هو حساب معامل الارتباط بطريقة كمية . وطريقة الحساب تتم بنفس المعادلات السابق استخدامها فى حساب الارتباط فى حالة البيانات غير المبوبة ، وذلك بعد إجراء بعض التعديل بما يتناسب مع طبيعة جدول التوزيع التكرارى المزدوج . إذ أن الأمر يتطلب أخذ التكرارت فى الحسبان ومراعاة أن س تمثل مراكز فئات س ، ص تمثل مراكز فــئات ص ، مجـ ك تمثل مجموع التكرارات .

وسوف نستعرض فيما يلى إحدى الطرق الخاصة بحساب معامل الارتباط فى حالة جدول التوزيع التكرارى المزدوج بشئ من الايجاز

طريقة الانحرافات باستخدام وسط فرضي :

وفقاً لهذه الطريقة فإننا نختار وسطاً فرضياً مناسباً لكل من المتغيرين س ، ص ويراعى في إختيار هذا الوسط الفرضي ، أن يكون عند مراكز إحدى الفئات ذات التكرارات الكبرى ، ويفضل بقدر الإمكان أن يكون في منتصف الجدول لكل من المتغيرين س ، ص ، وذلك حتى تكون انحرافات المراكز عنه أصغر ما يمكن .

وصيغة حساب معامل الارتباط بطريقة الانحرافات العادية في حالة كون البيانات مبوبة في صورة جدول توزيع تكراري مزدوج ، تكون معرفة كالآتي :

$$\text{ر} = \frac{\text{مجـ مجـ } ح_{س} ح_{ص} ك_{س ص} - (\text{مجـ ك})\left(\frac{\text{مجـ } ح_{س} ك_{س}}{\text{مجـ ك}}\right)\left(\frac{\text{مجـ } ح_{ص} ك_{ص}}{\text{مجـ ك}}\right)}{(\text{مجـ ك})\ ع_{س}\ ع_{ص}} \quad (٤)$$

وبالنسبة للانحرافين المعياريين الواردين بمقام المعادلة السابقة فإنهما يعرفان والحالة هذه كالآتي :

$$ع_{س} = \sqrt{\frac{\text{مجـ } ح_{س}^{2} ك_{س}}{\text{مجـ ك}} - \left(\frac{\text{مجـ } ح_{س} ك_{س}}{\text{مجـ ك}}\right)^{2}}$$

$$،\ ع_{ص} = \sqrt{\frac{\text{مجـ } ح_{ص}^{2} ك_{ص}}{\text{مجـ ك}} - \left(\frac{\text{مجـ } ح_{ص} ك_{ص}}{\text{مجـ ك}}\right)^{2}}$$

وفي صيغة صريحة فإن المعادلة السابقة تكون :

$$\text{ر} = \frac{\text{مجـ مجـ } ح_{س} ح_{ص} ك_{س ص} - (\text{مجـ ك})\left(\frac{\text{مجـ } ح_{س} ك_{س}}{\text{مجـ ك}}\right)\left(\frac{\text{مجـ } ح_{ص} ك_{ص}}{\text{مجـ ك}}\right)}{(\text{مجـ ك})\sqrt{\frac{\text{مجـ } ح_{س}^{2} ك_{س}}{\text{مجـ ك}} - \left(\frac{\text{مجـ } ح_{س} ك_{س}}{\text{مجـ ك}}\right)^{2}}\ \sqrt{\frac{\text{مجـ } ح_{ص}^{2} ك_{ص}}{\text{مجـ ك}} - \left(\frac{\text{مجـ } ح_{ص} ك_{ص}}{\text{مجـ ك}}\right)^{2}}} \quad (٥)$$

حيث :

$ح_س$: تمثل انحرافات مراكز فئات المتغير س عن الوسط الفرضى الخاص به .

$ح_ص$: تمثل انحرافات مراكز فئات المتغير ص عن الوسط الفرضى الخاص به .

$ك_{س ص}$ التكرارات المشتركة للمتغيرين س ، ص . وهى تلك التكرارات الواردة فى خلايا جدول التوزيع التكرارى المزدوج .

$مج ك_س = مج ك_ص = مج ك$ مجموع التكرارات :

هذا وحساب معامل الارتباط لبيانات جدول التوزيع التكرارى المزدوج الخاص بأوزان وأطوال ٥٠ شخصاً والوارد فى ص٢٢٣ يتطلب تكوين ثلاثة جداول :

الجدول الأول:

يمثل التوزيع التكرارى البسيط للمتغير س . ويتم تكوينه بأخذ فئات المتغير س الواردة فى الصف الأول من الجدول المزدوج والمجاميع الهامشية المناظرة لها والواردة فى الصف الأخير .

ثم نحسب من هذا الجدول الانحراف المعيارى للظاهرة س

الجدول الثانى :

يمثل التوزيع التكرارى البسيط للتغير ص ويتم تكوينه بأخذ فئات المتغير ص الواردة فى العمود الأول من الجدول المزدوج والمجاميع الهامشية المناظرة لها والواردة فى العمود الأخير .

ثم نحسب من هذا الجدول الانحراف المعيارى للتغير ص وبذلك نكون قد حصلنا على مقام معادلة الارتباط (٤) .

الجدول الثالث :

ويمثل جدول التوزيع التكرارى المزدوج وهو معطاة أصلا لنا .

ومن هذا الجدول نقوم بحساب بسط معامل الارتباط على الوجه الذى سنشرحه تفصيلا فى هذا الجزء .

(١) جدول التوزيع التكرارى البسيط للمتغير س

فئات س	$ك_{س}$	المراكز س	الانحراف عن ٧٢ $ح_{س}$	$ح_{س} \times ك_{س}$	$ح^{٢}_{س} \times ك_{س}$
٦٢ــ	٥	٦٤	ــ ٨	ــ ٤٠	٣٢٠
٦٦ــ	١٢	٦٨	ــ ٤	ــ ٤٨	١٩٢
٧٠ــ	١٨	٧٢	٠	٠	٠
٧٤ــ	١٣	٧٦	٤	٥٢	٢٠٨
٧٨ــ٨٢	٢	٨٠	٨	١٦	١٢٨
	٥٠			ــ ٢٠	٨٤٨

$$\text{الانحراف المعيارى} = ع_{س} = \sqrt{\frac{مجـ\, ح^{٢}_{س}\, ك_{س}}{مجـ\, ك} - \left(\frac{مجـ\, ح_{س}\, ك_{س}}{مجـ\, ك}\right)^{٢}}$$

$$= \sqrt{\frac{٨٤٨}{٥٠} - \left(\frac{-٢٠}{٥٠}\right)^{٢}}$$

$$= \sqrt{١٦{,}٩٦ - ٠{,}١٦}$$

$$= \sqrt{١٦{,}٨٠} = \underline{\underline{٤{,}١}}$$

(٢) جدول التوزيع التكرارى للمتغير ص

فئات ص	ك ص	المراكز ص	الانحراف عن ١٧٢,٥ ح ص	ح ص × ك ص	ح² ص × ك ص
١٦٠—	٤	١٦٢,٥	١٠—	٤٠—	٤٠٠
١٦٥—	١٢	١٦٧,٥	٥—	٦٠—	٣٠٠
١٧٠—	١٤	١٧٢,٥	٠	٠	٠
١٧٥—	١٠	١٧٧,٥	٥	٥٠	٢٥٠
١٨٠—	٧	١٨٢,٥	١٠	٧٠	٧٠٠
١٨٥—١٩٠	٣	١٨٧,٥	١٥	٤٥	٦٧٥
	٥٠			٦٥	٢٣٢٥

$$ع_{ص} = \sqrt{\frac{مج ح^2_{ص} ك_{ص}}{مج ك} - \left(\frac{مج ح_{ص} ك_{ص}}{مج ك}\right)^2}$$

$$= \sqrt{\frac{٢٣٢٥}{٥٠} - \left(\frac{٦٥}{٥٠}\right)^2}$$

$$= \sqrt{٤٦,٥٠ - ١,٦٩}$$

$$= \sqrt{٤٤,٨١} = ٦,٧$$

(٢) جدول التوزيع التكراري المزدوج للمتغيرين س ، ص معا

ح ص	فئات س / فئات ص	٦٢-	٦٦-	٧٠-	٧٤-	٧٨- ٨٢	مجموع حواصل الضرب مجـ ح س ح ص ك س ص
ح س		٨-	٤-	٠	٤	٨	
١٠-	١٦٠-	٢ [١٦٠]	٢ [٨٠]				٢٤٠
٥-	١٦٥-	٣ [١٢٠]	٦ [١٢٠]	٣ [صفر]			٢٤٠
٠	١٧٠-		٤ [صفر]	٨ [صفر]	٢ [صفر]		٠
٥	١٧٥-			٦ [صفر]	٤ [٨٠]		٨٠
١٠	١٨٠-			١ [صفر]	٥ [٢٠٠]	١ [٨٠]	٢٨٠
١٥	١٨٥- ١٩٠				٢ [١٢٠]	١ [١٢٠]	٢٤٠
	مجموع حواصل الضرب مجـ ح س ح ص ك س ص	٢٨٠	٢٠٠	٠	٤٠٠	٢٠٠	١٠٨٠

لحساب بسط معامل الارتباط المبينة (٤) من جدول التوزيع التكرارى المزدوج عن طريق استخدام وسط فرضى لكل من الظاهرين س ، ص أجرينا الآتى :

١ — اخترنا وسطاً فرضياً مناسباً للمتغير س عند س = ٧٢ . ثم حسبنا إنحراف كل مركز من مراكز فئات س عن هذا الوسط الفرضى ، ثم وضعنا هذه الانحرافات أعلى الجدول أمام الفئات الناظرة لها على التوالى .

وبالمثل إخترنا وسطاً فرضياً مناسباً للمتغير ص عند ص = ١٧٢٫٥ . ثم حسبنا إنحرافات كل مركز من مراكز فئات ص عند هذا الوسط الفرضى ، ثم وضعنا هذه الانحرافات الناتجة إلى اليمين من الجدول أما الفئات الناظرة لها على التوالى .

٢ — القيمة الموجودة داخل كل خلية والمحاطة بمستطيل أسفل التكرار عبارة عن حاصل ضرب التكرار الواقع فى هذه الخلية فى الانحرافين الناظرين لها أفقياً ورأسياً خارج الجدول . وهذا يعنى أن هذه القيمة عبارة عن حاصل ضرب ثلاثة قيم هى :

$$\text{ح}_{\text{س}} \times \text{ح}_{\text{ص}} \times \text{ك}_{\text{سص}}$$

فمثلا القيمة الموجودة فى الخلية الاولى من الصف الاول تحت التكرار ٢ وقيمتها ١٦٠ ناتجة من ضرب : — ١٠ × — ٨ × ٢ .

وبالمثل القيمة الموجودة فى الخلية الثانية من الصف الاول تحت التكرار ٢ وقيمتها ٨٠ ناتجة من ضرب : — ١٠ × — ٤ × ٢ .

وبالمثل القيمة الموجودة فى الخلية الأولى من الصف الثانى تحت التكرار ٣ وقيمتها ١٢٠ ناتجة من ضرب : — ٥ × — ٨ × ٢ .

وهكذا بالنسبة لسائر خلايا الجدول الأخرى .

٣ — نجمع حواصل الضرب السابقة بإشاراتها جمعاً جبرياً أفقياً وندون حاصل الجمع فى العمود الأخير المعنون بـ $\text{ح}_{\text{س}}\ \text{ح}_{\text{ص}}\ \text{ك}_{\text{سص}}$

وبالمثل نجمع حواصل الضرب السابقة بإشاراتها جمعاً جبرياً رأسياً وندون حاصل الجمع في الصف الأخير المعنون $\text{مجـ}\ ح_{س}\ ح_{ص}\ ك_{سص}$

٤ ـ مجموع الصف الأخير = مجموع العمود الأخير = $\text{مجـ}\ ح_{س}\ ح_{ص}\ ك_{سص}$

= ١٠٨٠

بالتعويض فى المعادلة (٤) نحصل على :

$$ر = \frac{\text{١٠٨٠} - \text{٥٠}\left(\frac{-\text{٢٠}}{\text{٥٠}}\right)\left(\frac{\text{٦٥}}{\text{٥٠}}\right)}{\text{٥٠} \times \text{٤,١} \times \text{٦,٧}}$$

$$= \frac{\text{١٠٨٠} + \text{٢٦}}{\text{١٣٧٣,٥}} = \frac{\text{١١٠٦}}{\text{١٣٧٣,٥}} = \text{٠,٨١}$$

وما يعنى وجود ارتباط طردى قوى بين المتغيرين س ، ص

ملخص المعادلات المستخدمة

طرق حساب معامل بيرسون للارتباط

البيانات غير المبوبة

١ - الطريقة المطولة باستخدام الوسط الحسابى للمتغيرين :

$$\text{ر} = \frac{\text{مجـ}(\text{س} - \bar{\text{س}})(\text{ص} - \bar{\text{ص}})}{\sqrt{\text{مجـ}(\text{س} - \bar{\text{س}})^2}\ \sqrt{\text{مجـ}(\text{ص} - \bar{\text{ص}})^2}}$$

٢ - الطريقة المباشرة باستخدام القيم الاصلية للمتغيرين

$$\text{ر} = \frac{\text{مجـ س ص} - \text{ن}\,\bar{\text{س}}\,\bar{\text{ص}}}{\sqrt{\text{مجـ س}^2 - \text{ن}\,\bar{\text{س}}^2}\ \sqrt{\text{مجـ ص}^2 - \text{ن}\,\bar{\text{ص}}^2}}$$

٣ - طريقة الانحرافات باستخدام وسط فرضى مناسب لكل من المتغيرين :

$$\text{ر} = \frac{\text{مجـ}\ \text{ح}_{\text{س}}\,\text{ح}_{\text{ص}} - \text{ن}\,\bar{\text{ح}}_{\text{س}}\,\bar{\text{ح}}_{\text{ص}}}{\sqrt{\text{مجـ}\ \text{ح}_{\text{س}}^2 - \text{ن}\,\bar{\text{ح}}_{\text{س}}^2}\ \sqrt{\text{مجـ}\ \text{ح}_{\text{ص}}^2 - \text{ن}\,\bar{\text{ح}}_{\text{ص}}^2}}$$

البيانات المبوبة

طريقة الانحرافات العادية باستخدام وسط فرضى :

$$\text{ر} = \frac{\text{مجـ مجـ}\ \text{ح}_{\text{س}}\,\text{ح}_{\text{ص}}\,\text{ك}_{\text{س ص}} - (\text{مجـ ك})\left(\frac{\text{مجـ}\ \text{ح}_{\text{س}}\,\text{ك}_{\text{س}}}{\text{مجـ ك}}\right)\left(\frac{\text{مجـ}\ \text{ح}_{\text{ص}}\,\text{ك}_{\text{ص}}}{\text{مجـ ك}}\right)}{(\text{مجـ ك})\sqrt{\frac{\text{مجـ}\ \text{ح}_{\text{س}}^2\,\text{ك}_{\text{س}}}{\text{مجـ ك}} - \left(\frac{\text{مجـ}\ \text{ح}_{\text{س}}\,\text{ك}_{\text{س}}}{\text{مجـ ك}}\right)^2}\ \sqrt{\frac{\text{مجـ}\ \text{ح}_{\text{ص}}^2\,\text{ك}_{\text{ص}}}{\text{مجـ ك}} - \left(\frac{\text{مجـ}\ \text{ح}_{\text{ص}}\,\text{ك}_{\text{ص}}}{\text{مجـ ك}}\right)^2}}$$

الباب الخامس

الانحدار الخطى

الانحـــدار الخطى البســـــيط

تمهـــــيد :

فى تحليل العلاقة بين متغيرين ، نهتم فى هذا الباب بدراسة الانحدار الخطى البسيط بين متغيرين س ، ص .

ومن الأهداف الرئيسية لدراسة الانحدار الخطى البسيط والتى يمكن تلخيصها فى الآتى :

أولاً : الحصول على معادلة رياضية للخط الذى يصف العلاقة الخطية المتوسطة the average linear relationship بين متغيرين اثنين فقط ، إحداهما المتغير التابع (ص) والآخر المتغير المستقل (س) .

وفى هذه الحالة يمكن استخدام هذه المعادلة فى التنبؤ أو تقدير قيم المتغير التابع بمعلومية القيم التى يأخذها المتغير المستقل .

ثانياً : الحصول على مقياس للخطأ الناشىء عن استخدام معادلة الانحدار كأساس للتقدير . والأمر يتطلب لتحقيق هذا الهدف حساب الخطأ المعيارى للتقدير the standand error of estimate والذى يقيس درجة تشتت قيم ص المشاهدة عن قيم صَ والمقدرة من معادلة خط الانحدار المحتسبة .

ثالثاً : الحصول على مقياس لتحديد نسبة ما يمكن تفسيره من التغير (التباين) فى المتغير التابع (ص) نتيجة للتغيرات التى تحدث فى المتغير المستقل (س) الداخل فى علاقة الانحــدار .

والأمر يتطلب لتحقيق هذا الهدف حساب معامل التحديد the coefficient of determination والذى يقيس مدى جودة توفيق goodness - of - fit خط الانحدار للبيانات موضوع الدراسة .

وفيما يختص بالخصائص الاحصائية لمقدرات نموذج خط الانحدار وكذلك اختبارات الفروض للمعالم الفردية وللنموذج ككل ، فإنها جميعها تخرج عن نطاق هذا المقرر وازاء ذلك فإننا سوف لا نتعرض لها فى هذا الباب .

نموذج الانحدار الخطى البسيط :

Simple Linear Regression Model

يعتبر نموذج الانحدار الخطى البسيط من أكثر نماذج الانحدار بساطة وسهولة واستخداما فى التقدير والتحليل الاحصائى للعلاقة بين متغيرين اثنين فقط : أحداهما متغيراً تابعاً ويرمز له بالرمز ص والآخر متغيراً مستقلاً ويرمز له بالرمز س . وهذا النموذج يمكن وصفه بالمعادلة الآتية (١) :

ص = α + β س + ε (١)

حيث :

ص : تعبر عن المتغير التابع

س : تعبر عن المتغير المستقل

α ، β : تعبر عن معالم أو ثوابت المجتمع المجهولة .

ε : تعبر عن حد الخطأ .

وهناك افتراضات هامه ، تتطلب الدراسة والتحليل توافرها فى هذا النموذج وسوف نتناولها باختصار فيما يلى :

أولاً : أن قيمة المتغير التابع ص تتوقف إلى حد ما على القيمة التى يأخذها المتغير المستقل س . وهذا يعنى أن يكون المتغير التابع متغيراً عشوائياً ، بينما المتغير المستقل يكون مثبتاً ومحدداً عند مستويات معينة تخضع لاختيار وتحكم الباحث .

ثانياً : عند كل قيمة من قيم س ، تأخذ ص مجموعة من القيم المختلفة . وأن توزيع هذه القيم الصادية يكون توزيعاً معتاداً . وأن متوسطات هذه التوزيعات المعتادة لـ ص تكون جميعها واقعة على خط الانحدار ، كما وأن أخطائها المعيارية تكون متساوية .

---

(١) يمكن تبرير ادخال حد الخطأ ε فى نموذج الانحدار الخطى إلى عدة عوامل منها : (أ) اهمال ادخال بعض متغيرات مستقلة أخرى فى النموذج لكون تأثيرها ضئيل على المتغير التابع (ب) الأخطاء الناتجة عن قياس المتغير التابع (جـ) أخطاء المعاينة .

ثالثاً : أن العلاقة المتوسطة بين المتغيرين س ، ص فى المجتمع ، يمكن وصفها بكفاية adequately بالمعادلة الخطية والتى يمكن التعبير عنها هندسياً بخط مستقيم . وكل نقطة على هذا الخط المستقيم توضح القيمة المتوسطة للمتغير التابع ص عندما تكون قيمة المتغير المستقل س مثبته عند قيمة معينة .

هذا وعندما تكون قيمة س = صفراً ، فإن القيمة المتوسطة للمتغير التابع ص تكون مساوية لقيمة الثابت α ، وهى تمثل نقطة تقاطع خط الانحدار مع المحور الصادى .

وميل هذا الخط المستقيم أو كما يعبر عنه بالثابت β (معامل انحدار ص/ س) فى المعادلة الخطية ، فيوضح التغير المتوسط الذى يحدث فى المتغير التابع ص لكل وحدة تغير واحدة فى المتغير المستقل س . وعندما يميل هذا الخط إلى أعلى جهة اليمين فإن اشارة β تكون موجبة ، وهذا يعنى أن العلاقة بين المتغيرين علاقة طردية . وبالعكس من ذلك إذا كان هذا الخط يميل إلى أسفل جهة اليمين فإن اشارة β تكون سالبة ، وهذا يعنى أن العلاقة بين المتغيرين علاقة عكسية . وعلى ذلك نجـد أن اشارة β توضح طبيعة العلاقة بين المتغيرين موضوع الدراسة .

هذا ولما كان الخط المستقيم يمكن تعريفه تماما بمعلومـية الجزء المقـطوع من المحور الصادى وميله ، ولذلك فإن مهمة تقدير خط انحدار المجتمع ليس إلا تقديراً لقيمتى α ، β .

رابعاً : أن تكون قيم ص مستقلة احصائياً ، وهذا يعنى أنه عند اختيار العينة تكون قيم ص عند أى قيمة لـ س لا تتوقف على قيم أخرى لـ ص لأى قيمة أخرى لـ س .

خامساً : أن تكون حدود الخطأ موزعة توزيعاً معتاداً بمتوسط صفر وتباين ثابت σ^2 .

كما أنه يفترض أن تكون حدود الخطأ مستقلة عن بعضها البعض ، ومستقلة أيضاً عن السينات

تقدير معلمات خط الانحدار α ، β بطريقة المربعات الصغرى :

فى تحليل الانحدار الخطى البسيط تكون مهمتنا الأولى هى تقدير قيمة معلمتى انحدار المجتمع المجهولتين α ، β فى معادلة خط الانحدار (١) على أساس ن من أزواج المشاهدات فى العينة .

وإذا رمزنا لتقديرات α ، β من العينة بالرموز أ ، ب على التوالى ، فإن معادلة خط الانحدار فى العينة the sample regression equation تكون على الصورة الآتية :

$$\hat{ص} = أ + ب\ س \qquad (٢)$$

وهناك عدة طرق لتقدير قيمتى α ، β أى حساب أ ، ب من العينة . إلا أن أفضل هذه الطرق هى طريقة المربعات الصغرى[١] Least squares method والتى تتميز بالخاصيتين الآتيتين :

الأولى : مجـ (ص − ص̂) = صفراً

بمعنى أن المجموع الجبرى للانحرافات الرأسية لقيم المتغير ص الفعلية عن خط الانحدار تساوى صفراً . وبعبارة أخرى أن مجموع انحرافات قيم ص الفعلية عن قيم ص̂ (والمقدرة من معادلة خط الانحدار) يساوى صفراً .

الثانية : مجـ $(ص - \hat{ص})^{2}$ = أصغر ما يمكن .

بمعنى أن مجموع مربعات الانحرافات الرأسية لقيم ص الفعلية عن خط الانحدار تكون أصغر ما يمكن ، وبعبارة أخرى أن مجموع مربعات انحرافات قيم ص الفعلية عن قيم ص̂ (والمقدرة من معادلة خط الانحدار) يكون أصغر مايمكن الحصول عليه من استخدام أى خط مستقيم آخر .

---

(١) تقدير ثوابت الانحدار باستخدام هذه الطريقة تتميز بعدم التحيز وأن تبايناتها تكون أصغر مايمكن .

هذا ويمكن إيجاز ما تقدم فى أن طريقة المربعات الصغرى تمكننا من تقدير معلمات انحدار المجتمع α ، β أى حساب أ ، ب من بيانات العينة بحيث تكون خاضعة للشرط[1] :

مجـ خ^{٢} = مجـ (ص − صَ)^{٢} = مجـ (ص − أ − ب س)^{٢} = أصغر مايمكن (٣)

ومن المعلوم أن هذا الشرط يتحقق إذا حسبت كل من أ ، ب من حل المعادلتين الطبيعيتين الموضحتين فيما يلى the two normal equations واللتين يمكن اشتقاقهما بسهولة من المعادلة (٢) كالآتى :

أولاً : بضرب المعادلة (٢) فى معامل المجهول الأول أ والذى يساوى ١ . ثم بالجمع بالنسبة للعدد الكلى ن من المشاهدات فى العينة نحصل على المعادلة الطبيعية الأولى على الوجه الآتى :

مجـ$_{١}^{ن}$ { ١ (ص = أ + ب س) } = مجـ$_{١}^{ن}$ { ص = أ + ب س }

أو ببساطة تكون المعادلة الطبيعية الأولى هى :

مجـ ص = ن أ + ب مجـ س

ثانياً : بضرب المعادلة (٢) فى معامل المجهول الثانى ب والذى يساوى س . ثم بالجمع بالنسبة للعدد الكلى ن من المشاهدات فى العينة نحصل على المعادلة الطبيعية الثانية على الوجه الآتى :

مجـ$_{١}^{ن}$ { س (ص = أ + ب س) } = مجـ$_{١}^{ن}$ { س ص = أ س + ب س^{٢} }

أو ببساطة تكون المعادلة الطبيعية الثانية هى :

مجـ س ص = أ مجـ س + ب مجـ س^{٢}

---

(١) من الواضح أن الخاصية الأولى تتحقق تلقائيا عن تحقيق هذا الشرط

وعلى ذلك تكون المعادلتين الطبيعيتين هما (١) :

مجـ ص = ن أ + ب مجـ س (٤)

مجـ س ص = أ مجـ س + ب مجـ س٢ (٥)

ويحل هاتين المعادلتين مع بعضهما جبرياً باستخدام طريقة الحذف the method elimination نحصل على قيم كل من أ ، ب .

طريقة أخرى لحساب ثوابت الانحدار :

بدلا من استخدام طريقة الحذف لإيجاد قيم كل من أ ، ب فإنه يمكننا بالتعويض فى المعادلتين (٦) ، (٩) الحصول على قيمتى هذين الثابتين مباشرة وهاتين المعادلتين يمكن اشتقاقهما بسهولة من المعادلتين الطبيعتين (٤) ، (٥) على الوجه الآتى :

أولاً : حساب قيمة الثابت أ :

من المعادلة الطبيعية الأولى رقم (٤) نجد أن :

$$أ = \frac{مجـ ص}{ن} - ب \frac{مجـ س}{ن}$$

$$\therefore أ = \bar{ص} - ب \bar{س} \quad (٦)$$

---

(١) يمكن باستخدام أسلوب التفاضل الجزئى للحصول على نفس هاتين المعادلتين الطبيعيتين وذلك بتفاضل مجـ غ٢ = مجـ (ص – أ – ب س)٢ جزئياً مرة بالنسبة إلى أ ثم مرة أخرى بالنسبة إلى ب ثم مساواة الناتج بالصفر . وعلى ذلك فإن :

$$\frac{6 مجـ غ^٢}{16} = - ٢ مجـ (ص - أ - ب س) صفر$$

$$\therefore مجـ (ص - أ - ب س) = صفر$$

أو ببساطة مجـ ص = ن أ + ب مجـ س وهى نفس المعادلة الطبيعية الأولى رقم (٤)

$$\frac{6 مجـ غ^٢}{6ب} = - ٢ مجـ س (ص - أ - ب س) صفر$$

$$= - ٢ مجـ (س ص - أ س - ب س^٢) = صفر$$

$$\therefore مجـ (س ص - أ س - ب س^٢) = صفر$$

أو ببساطة مجـ س ص = أ مجـ س + ب مجـ س٢ وهى نفس المعادلة الطبيعية الثانية رقم (٥)

<u>ثانياً : حساب قيمة الثابت ب :</u>

• باستخدام طريقة المحددات لحساب قيمة ب من المعادلتين الطبيعيتين رقمى (٤) ، (٥) نجد أن :

$$ب = \frac{\begin{vmatrix} ن & مجـ ص \\ مجـ س & مجـ س ص \end{vmatrix}}{\begin{vmatrix} ن & مجـ س \\ مجـ س & مجـ س^2 \end{vmatrix}} = \frac{ن \text{ مجـ س ص} - (\text{مجـ س})(\text{مجـ ص})}{ن \text{ مجـ س}^2 - (\text{مجـ س})^2} \quad (٧)$$

بقسمة البسط والمقام من المعادلة السابقة على ن ينتج أن :

$$ب = \frac{\text{مجـ س ص} - \frac{(\text{مجـ س})(\text{مجـ ص})}{ن}}{\text{مجـ س}^2 - \frac{(\text{مجـ س})^2}{ن}} \quad (٨)$$

$$\text{أو } ب = \frac{\text{مجـ س ص} - ن\, \bar{س}\, \bar{ص}}{\text{مجـ س}^2 - ن\, \bar{س}^2} \quad (٩)$$

طريقة انحرافات المتغيرات عن أوساطها الحسابية :

يمكن حساب ثوابت الانحدار أ ، ب بطريقة مبسطة إذا كانت البيانات الرقمية للمتغيرات كبيرة نسبياً مما يترتب عليه أن حساب المربعات وحواصل الضرب تتطلب بعض الوقت والجهد .

هذا ، وإذا ما أخذنا انحرافات قيم كل متغير من المتغيرات الداخلة فى علاقة الانحدار عن وسطه الحسابى . فإنه طبقاً للخصائص التى يتميز بها الوسط الحسابى وهى أن " مجموع انحرافات قيم المتغير عن وسطه الحسابى يساوى صفراً " سوف يترتب على ذلك أن :

$$\text{مجـ}(س - \bar{س}) = \text{صفراً}$$

$$، \text{مجـ}(ص - \bar{ص}) = \text{صفراً}$$

وعلى ذلك فإنه إذا استبدلنا قيم كل متغير بانحرافاته عن الوسط الحسابى ، فإننا نكون فى هذه الحالة قد نقلنا نقطة الأصل بالنسبة للمعادلتين الطبيعيتين (٤) ، (٥) من النقطة (٠ ، ٠) إلى النقطة ($\bar{\text{س}}$ ، $\bar{\text{ص}}$) . ويترتب على ذلك تخفيض عدد المعادلات الطبيعية من معادلتين إلى معادلة واحدة بدلالة انحرافات متغيراتها عن أوساطها الحسابية وهى :

$$\text{مجـ}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}}) = \text{ب}\ \text{مجـ}(\text{س}-\bar{\text{س}})^{2}$$

وتكون قيمة ب كالآتى :

$$\text{ب} = \frac{\text{مجـ}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}})}{\text{مجـ}(\text{س}-\bar{\text{س}})^{2}}$$

هذا ، ويلاحظ أننا لو قسمنا البسط والمقام فى المعادلة السابقة على (ن − ١) فإن المعادلة (١٠) تصير :

$$\text{ب} = \frac{\text{مجـ}(\text{س}-\bar{\text{س}})(\text{ص}-\bar{\text{ص}})/(\text{ن}-1)}{\text{مجـ}(\text{س}-\bar{\text{س}})^{2}/(\text{ن}-1)}$$

ويعرف البسط على أنه تغاير العينة sampie covariance للمتغيرين س ، ص . أما المقام فإنه يعرف على أنه تباين العينة بالنسبة للمتغير س . وعلى ذلك فإن المعادلة السابقة تعنى أن :

$$\text{ب} = \frac{\text{ع}_{\text{س ص}}}{\text{ع}^{2}_{\text{س}}} = \frac{\text{تغاير س ، ص}}{\text{تباين س}} \qquad (١١)$$

أما قيمة الثابت أ فتحسب كالمعتاد من المعادلة (٦) : $\text{أ} = \bar{\text{ص}} - \text{ب}\ \bar{\text{س}}$

مثال (١) :

الجدول الآتى يوضح المبالغ المنصرفة على بعض السلع الاستهلاكية بمئات الجنيهات (ص) فى السنة بالنسبة إلى الدخل السنوى بآلاف الجنيهات (س) وذلك من عينة مكونه من عشرة أسر اختيرت بطريقة عشوائية :

ص(الاستهلاك)	١٩	٢١	٢٢	٦	١٤	١٣	٩	١٦	٢٢	١٨
س (الدخل)	٦	٦	٧	١	٣	٥	٣	٥	٩	٥

والمطلوب حساب معادلة انحدار ص/ س .

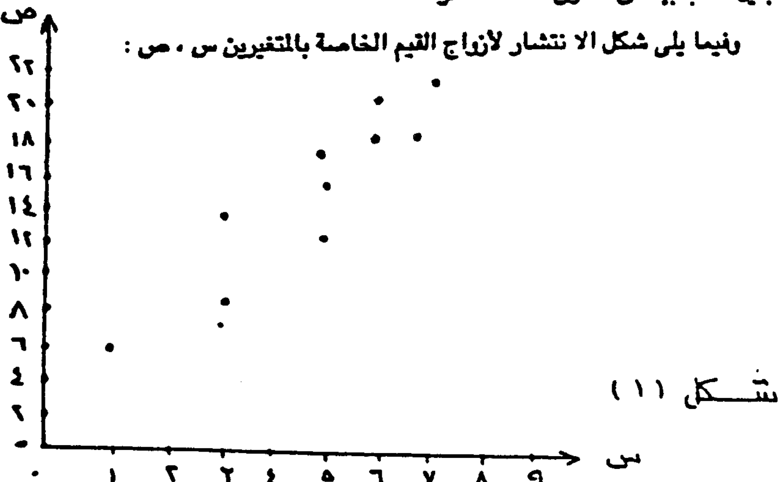
الحـل :

الخـطوة الأولى فى تحليل الانحدار هى تكوين نقط شكل الانتشار لأزواج القيم الواردة بالعينة بحيث يخصـص المحور الأفقى لتمثيل قيم المتغير المستقل وهو فى حالتنا هذه : الدخل (س) ، والمحور الرأسى لتمثيل قيم المتغير التابع وهو فى حالتنا هذه : الاستهلاك (ص) .

والخطوة التالية هى فحص نقط شكل الانتشار للوقوف عما إذا كان المتغير التابع يعتمد إلى حد ما على المتغير المستقل ، وعما كانت العلاقة المتوسطة بينهما يمكن تفسيرها بخط مستقيم يمر بين هذه النقط .

هذا ، وإذا ما أوضح الفحص النظرى أن تحليل الانحدار الخطى البسيط يمكن أن يصف بيانات العينة بكفاية adequately فإن المشكلة التى تصادفنا الآن هو توفيق خط مستقيم يتوسط نقط شكل الانتشار بحيث يمثلها أفضل تمثيل ويتم ذلك بحساب الثابتين أ ، ب بإحدى الطرق سالفة الذكر .

وفيما يلى شكل الا نتشار لأزواج القيم الخاصة بالمتغيرين س ، ص :



شكل (١)

ولحساب قيمتى أ ، ب ننشىء الجدول الآتى هذا مع ملاحظة أن عمود ص^{٢} قد أضيف للاستفادة منه مستقبلا فى حساب الخطأ المعيارى للتقدير .

س	ص	س ص	س^{٢}	ص^{٢}
٦	١٩	١١٤	٣٦	٣٦١
٦	٢١	١٢٦	٣٦	٤٤١
٧	٢٢	١٥٤	٤٩	٤٨٤
١	٦	٦	١	٣٦
٣	١٤	٤٢	٩	١٩٦
٥	١٣	٦٥	٢٥	١٦٩
٣	٩	٢٧	٩	٨١
٥	١٦	٨٠	٢٥	٢٥٦
٩	٢٢	١٩٨	٨١	٤٨٤
٥	١٨	٩٠	٢٥	٣٢٤
٥٠	١٦٠	٩٠٢	٢٩٦	٢٨٣٢

سبق أن ذكرنا أن معادلة خط انحدار ص/ س هى :

$$\hat{ص} = أ + ب س$$

وأن هناك عدة طرق لحساب قيم أ ، ب :

أولاً : طريقة الحذف :

بالتعويض بالقيم الواردة بالجدول أعلاه فى المعادلتين الطبيعيتين الآتيتين :

مجـ ص = ن أ + ب مجـ س

مجـ س ص = أ مجـ س + ب مجـ س

ينتج أن :

١٦٠ = ١٠أ + ٥٠ ب

٩٠٢ = ٥٠أ + ٢٩٦ ب

ثم نقوم بحل هاتين المعادلتين مع بعضهما كالآتى :

المعادلة الثانية كما هى : ٩٠٢ = ٥٠أ + ٢٩٦ ب
المعادلة الأولى × ٥ : ٨٠٠ = ٥٠أ + ٢٥٠ ب } بالطرح

١٠٢ = ٤٦ ب

$\therefore$ ب = $\frac{١٠٢}{٤٦}$ = **٢,٢١٧**

وبالتعويض بقيمة ب السابقة فى المعادلة الأولى ينتج أن :

١٦٠ = ١٠أ + ٥٠ (٢,٢١٧)

$\therefore$ ١٠أ = ١٦٠ − ١١٠,٨٥ = ٤٩,١٥

$\therefore$ أ = $\frac{٤٩,١٥}{١٠}$ = **٤,٩١٥**

وتكون معادلة خط انحدار ص/ س كالآتى :

$\hat{ص}$ = ٤,٩١٥ + ٢,٢١٧ س

ثانياً : الطريقة المباشرة :

من المعلوم من المعادلة (٩) أن :

ب = $\frac{\text{مجـ س ص} - \text{ن}\ \bar{س}\ \bar{ص}}{\text{مجـ س}^{٢} - \text{ن}\ \bar{س}^{٢}}$

= $\frac{٩٠٢ - ١٠(٥)(١٦)}{٢٩٦ - ١٠(٥)^{٢}}$ = $\frac{٩٠٢ - ٨٠٠}{٢٩٦ - ٢٥٠}$ = $\frac{١٠٢}{٤٦}$ = **٢,٢١٧**

وبالتعويض بقيمة ب فى المعادلة (٦) لإيجاد أ حيث :

أ = $\bar{ص}$ − ب $\bar{س}$

= ١٦ − (٢,٢١٧)(٥) = ١٦,٠٠٠ − ١١,٠٨٥ = **٤,٩١٥**

معادلة خط انحدار ص/ س تكون :

ص = ٤,٩١٥ + ٢,٢١٧ س

ومما تجدر الاشارة إليه أن قيمة أ - وهى تقدير لقيمة α - تنحصر أهميتها فقط فى الناحية الرياضية ، إذ أنها مع قيمة ب تحدد موضع خط الانحدار فى الرسم البيانى .

وبالعكس من ذلك فإن قيمة ب - وهى تقدير لقيمة β - لها أهمية ودلالة تطبيقية . فنجد فى مثالنا هذا أن قيمتها موجبة وتساوى ٢,٢١٧ وهى بذلك تشير إلى أنه كلما زاد دخل الأسرة بمقدار وحدة واحدة فى السنة (أى بمقدار الف جنيه فى السنة) فإن تقدير المبالغ المنصرفة على السلع الاستهلاكية سوف يزيد فى المتوسط بمقدار ٢,٢١٧ وحدة من مئات الجنيهات ، أى بما تساوى ٢٢١,٧ جنيهاً فى السنة .

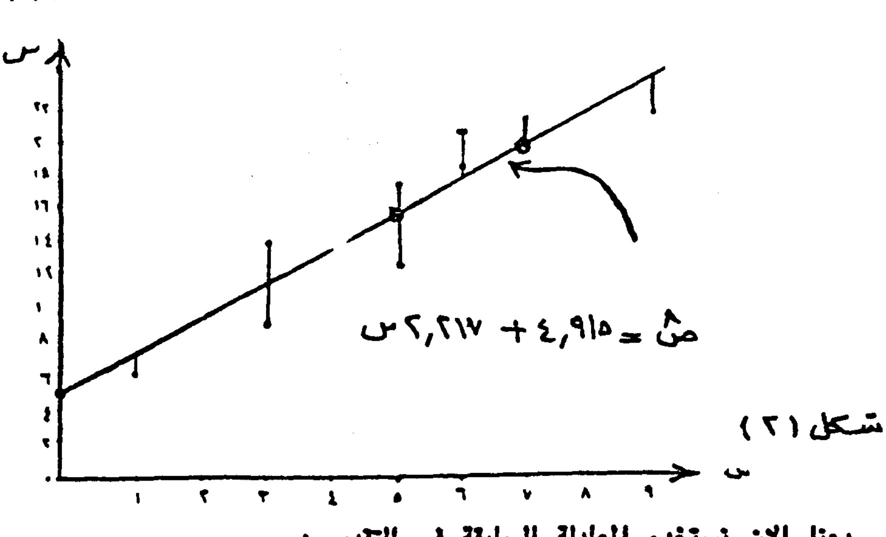
دعنا الآن نستخدم المعادلة السابقة فى رسم خط الانحدار

من المعلوم أنه يمكن رسم الخط المستقيم الممثل للدالة الخطية بواسطة نقطتين عليه إلا أنه من المفضل دائماً أخذ نقطة ثالثة للتحقق من دقة النقطتين الأخريتين .

نعطى س قيماً عددية ولتكن ٠ ، ٥ ، ٧ ثم نعوض بها فى معادلة الإنحدار وندون النتائج فى الجدول الآتى :

س	٠	٥	٧
ص	٤,٩	١٦	٢٠,٤

ثم نمثل كل زوج من هذه الأزواج بنقطة . ثم نوصل هذه النقط نحصل بذلك على خط الانحدار كما هو فى الشكل (٢) الآتى :



شكل (٢)

دعنا الان نستخدم المعادلة السابقة فى التقدير :

إذا افترضنا أننا أردنا تقدير المبالغ المنصرفة على السلع الاستهلاكية لعائلة مسحوبة من نفس المجتمع المسحوبة منه هذه العينة دخلها السنوى ٤٠٠٠ جنيه .

فى هذه الحالة نعوض عن س = ٤ وحدات فى معادلة الانحدار المشار إليها بعاليه فنجد أن : ص^ = ٤,٩١٥ + (٢,٢١٧) (٤)

= ١٢,٧٨٣ وحدة (من مئات الجنيهات) = ١٢٧٨,٣ جنيهاً .

وهذا يعنى أن تقدير المبالغ المنصرفة على السلع الاستهلاكية للعائلة ذات الدخل السنوى ٤٠٠٠ جنيهاً هو ١٢٧٨,٣ جنيهاً فى السنة .

تقدير التباين والخطأ المعيارى لخط الانحدار

معادلة خط الانحدار والمحسوبة من بيانات العينة وهى :

ص^ = أ + ب س

تستخدم للتقدير أو للتنبؤ بقيمة المتغير التابع ص عند قيمة محددة لـ س . ولكن المشكلة التى تواجهنا الآن هى أنه إلى أى مدى يمكن الاعتماد على هذه المعادلة

كأساس للتنبؤ أو للتقدير ؟ والإجابة على هذا السؤال يتوقف على مدى تقارب أو تباعد قيم ص الفعلية من القيم المقدرة (صً) من معادلة الانحدار أو خط الانحدار عند قيم محددة لـ س .

هذا ويمكننا الحصول على صورة مرئية عن طبيعة انتشار أو تشتت هذه القيم وذلك برسم خط الانحدار الممثل لمعادلة الانحدار سالفة الذكر والذى يمر خلال نقط شكل الانتشار لقيم ص الفعلية . ثم توصيل النقط التى فوق وتحت هذا الخط بخطوط رأسية كما هو فى الشكل (٢) بالصفحة السابقة والخاص ببيانات المثال رقم (١) .

ومن المعلوم أنه كلما قربت نقط شكل الانتشار من خط الانحدار كلما كانت معادلة الانحدار أكثر دقة فى تمثيل العلاقة بين المتغيرات الداخلة فى علاقة الانحدار وبالتالى فإنه يمكن الاعتماد عليها كأساس للتنبؤ أو للتقدير بقيم المتغير التابع ص عند قيم معينة للمتغير المستقل س .

هذا ، وبفحص الشكل البيانى (٢) الخاص ببيانات العينة نجد أن خط الانحدار الممثل لمعادلة الانحدار صً = ٤,٩١٥ + ٢,٢١٧ س تصف بكفاية العلاقة بين المتغيرين س ، ص موضع الدراسة .

ومن الأهمية بمكان هو الحصول على مقياس عددى a numerical measure لقياس مدى انتشار أو تشتت القيم الفعلية للمتغير التابع ص حول خط الانحدار وبحيث يمكن استخدامه كمؤشر لأخطاء التقدير the error of estimation وبعبارة أخرى نريد مقياساً لتقدير تباين الانحدار فى المجتمع(١) (أى $\sigma^2_{ص/س}$) أو لتقدير الخطأ المعيارى للانحدار فى المجتمع ($\sigma_{ص/س}$) وكالمعتاد سوف نعتمد على بيانات العينة لعمل هذه التقديرات .

---

(١) معادلة تباين الانحدار الخطى البسيط فى المجتمع هى :

$$\sigma^2_{ص/س} = \frac{مجـ (ص - \mu_{ص/س})^2}{ن}$$

وطبقاً للتعريف العام للتباين فإن تباين الانحدار من العينة sample variance of the regression يعرف كالآتى :

$$\text{ع}^{\text{٢}}_{\text{ص/س}} = \frac{\text{مجـ}(\text{ص} - \hat{\text{ص}})^{\text{٢}}}{\text{ن} - \text{٢}} \qquad (\text{١٢})$$

والمقام هنا (ن – ٢) يمثل درجات الحرية degrees of freedom والدرجتين المفقودتين جاءت نتيجة تقدير α ، β من بيانات العينة بحساب أ ، ب .

والجذر التربيعى الموجب للمعادلة (١٢) هو تقدير الخطأ المعيارى للانحدار فى المجتمع أى للخطأ المعيارى للمتغير التابع ص عند قيم محددة للمتغير المستقل س . وهو بذلك يقيس مدى قرب closeness تقديرات المتغير التابع المحسوبة من معادلة الانحدار من القيم الفعلية لهذا المتغير ويرمز له بالرمز $\text{ع}_{\text{ص/س}}$.

وعلى ذلك فإن :

$$\text{ع}_{\text{ص/س}} = \sqrt{\frac{\text{مجـ}(\text{ص} - \hat{\text{ص}})^{\text{٢}}}{\text{ن} - \text{٢}}} \qquad (\text{١٣})$$

والمقياس السابق يعرف عادة بالخطأ المعيارى للتقدير the standard error of estimate وكلما كانت قيمة هذا المقياس العددية صغيرة كلما كانت القيم الفعلية للمتغير التابع ص قريبة من خط الانحدار وبالتالى تكون معادلة الانحدار المتخذة أساساً للتقدير أو للتنبؤ أكثر دقة (١) .

---

(١) ويفيدنا الخطأ المعيارى للتقدير فى أنه إذا استخدمت معادلة خط الانحدار ص/ س فى تقدير قيم ص ، فإننا نستطيع القول أن حوالى ٦٨,٢٦ ٪ من قيم ص والمرتبطة بقيمة محددة من قيم س تقع فى المساحة المحصورة بين (+ $\text{ع}_{\text{ص/س}}$ ، – $\text{ع}_{\text{ص/س}}$) حول خط الانحدار ، وأن حوالى ٩٥,٤٥ ٪ من قيم ص سوف تقع فى المساحة المحصورة بين (+ ٢$\text{ع}_{\text{ص/س}}$ ، – ٢$\text{ع}_{\text{ص/س}}$) حول خط الانحدار ، وأن حوالى ٩٩,٧٣ ٪ من قيم ص سوف تقع فى المساحة المحصورة بين (+ ٣$\text{ع}_{\text{ص/س}}$ ، – ٣$\text{ع}_{\text{ص/س}}$) حول خط الانحدار ، وذلك بافتراض أن توزيع ص يتبع التوزيع المعتاد

خطوات حساب الخطأ المعيارى للتقدير :

على ضوء ما تقدم فإن خطوات حساب الخطأ المعيارى للتقدير لمتغيرين س ، ص حيث : ص ترمز للمتغير التابع ، س ترمز للمتغير المستقل تكون كالآتى :

أولاً : نحسب معادلة خط الانحدار ص/ س طبقاً لطريقة المربعات الصغرى .

ثانياً : نحسب تقديرات قيم المتغير التابع وذلك بالتعويض بقيم المتغير المستقل المتناظرة فى معادلة خط الانحدار المحسوبة فى الخطوة السابقة . ويرمز لهذه التقديرات بالرمز صَ .

ثالثاً : نحسب الفروق (الأخطاء) بين قيم المتغير التابع الفعلية ص وقيم المتغير التابع المقدرة من معادلة خط الانحدار أى صَ . وسوف نرمز لها بالرمز خ .

وطبقاً لطريقة المربعات الصغرى فإن مجموع هذه الفروق (الأخطاء) لا بد وأن تكون مساوية للصفر ، بمعنى أن :

$$\text{مجـ خ} = \text{مجـ}(\text{ص} - \text{صَ}) = \text{صفراً}$$

رابعاً : نربع قيم الأخطاء الواردة فى عمود خ ثم نجمع هذا العمود نحصل بذلك على

$$\text{مجـ خ}^2 = \text{مجـ}(\text{ص} - \text{صَ})^2$$

خامساً : نحسب تباين الانحدار من المعادلة (١٢) السابق ذكرها وهى :

$$\text{ع}^2_{\text{ص/س}} = \frac{\text{مجـ}(\text{ص} - \text{صَ})^2}{\text{ن} - 2}$$

سادساً : بأخذ الجذر التربيعى للتباين السابق نحصل بذلك على الخطأ المعيارى للتقدير . وهذا يعنى أن :

$$\text{ع}_{\text{ص/س}} = \sqrt{\frac{\text{مجـ}(\text{ص} - \text{صَ})^2}{\text{ن} - 2}}$$

مثال (٢) :

احســب الخطأ المعــيارى لخط انحدار ص/ س من بيانات العينة الواردة فى المثال(١).

الحــل :

سبق أن حسبنا معادلة الانحدار الخطى البسيط لبيانات هذه العينة ووجدناها كالآتى : صً = ٤,٩١٥ + ٢,٢١٧ س

وفيما يلى تقديرات قيم المتغيرات التابع المحسوبة من المعادلة السابقة عند كل قيمة من قيم المتغير المستقل س :

ص|٦ = ٤,٩١٥ + (٢,٢١٧)(٦) = ١٨,٢١٧

ص|٦ = ٤,٩١٥ + (٢,٢١٧)(٦) = ١٨,٢١٧

ص|٧ = ٤,٩١٥ + (٢,٢١٧)(٧) = ٢٠,٤٣٤

ص|١ = ٤,٩١٥ + (٢,٢١٧)(١) = ٧,١٣٢

ص|٣ = ٤,٩١٥ + (٢,٢١٧)(٣) = ١١,٥٦٦

ص|٥ = ٤,٩١٥ + (٢,٢١٧)(٥) = ١٦,٠٠٠

ص|٣ = ٤,٩١٥ + (٢,٢١٧)(٣) = ١١,٥٦٦

ص|٥ = ٤,٩١٥ + (٢,٢١٧)(٥) = ١٦,٠٠٠

ص|٩ = ٤,٩١٥ + (٢,٢١٧)(٩) = ٢٤,٨٦٨

ص|٥ = ٤,٩١٥ + (٢,٢١٧)(٥) = ١٦,٠٠٠

مجـ صً = ١٦٠,٠٠٠

وهذا يعنى أن : مجـ ص = مجـ صً

وبعبارة أخرى أن : مجـ (ص − صً) = صفراً

وفيما يلى جدول حساب الخطأ المعيارى للتقدير

القيم المشاهدة س	القيم المشاهدة ص	القيم المقدرة صَ	الأخطاء (ص - صَ)	مربع الأخطاء (ص - صَ)²
٦	١٩	١٨,٢١٧	٠,٧٨٣	٠,٦١٣٠٨٩
٦	٢١	١٨,٢١٧	٢,٧٨٣	٧,٧٤٥٠٨٩
٧	٢٢	٢٠,٤٣٤	١,٥٦٦	٢,٤٥٢٣٥٦
١	٦	٧,١٣٢	١,١٣٢-	١,٢٨١٤٢٤
٣	١٤	١١,٥٦٦	٢,٤٣٤	٥,٩٢٤٣٥٦
٥	١٣	١٦,٠٠٠	٣,٠٠٠-	٩,٠٠٠٠٠٠
٣	٩	١١,٥٦٦	٢,٥٦٦-	٦,٥٨٤٣٥٦
٥	١٦	١٦,٠٠٠	-	-
٩	٢٢	٢٤,٨٦٨	٢,٨٦٨-	٨,٢٢٥٤٢٤
٥	١٨	١٦,٠٠٠	٢,٠٠٠	٤,٠٠٠٠٠٠
٥٠	١٦٠	١٦٠,٠٠٠	٩,٥٦٦ ٩,٥٦٦- صفر	٤٥,٨٢٦٠٩٤

$$\text{ع}^{2}_{\text{ص/س}} = \frac{\text{٤٥,٨٢٦٠٩٤}}{\text{١٠} - \text{٢}} = \frac{\text{٤٥,٨٢٦٠٩٤}}{\text{٨}} = \text{٥,٧٢٨٢٦٢}$$

$$\text{ع}_{\text{ص/س}} = \sqrt{\text{٥,٧٢٨٢٦٢}} = \underline{\underline{\text{٢,٣٩٣}}} \quad (\text{الخطأ المعيارى للتقدير})$$

حساب الخطأ المعيارى للتقدير [1] بدلالة القيم المعطاة :

من الواضح أن حساب تباين خط انحدار ص/ س والذى جذره التربيعى الموجب عبارة عن الخطأ المعيارى للتقدير باستخدام المعادلة (١٢) يتطلب وقتاً وجهداً .

هذا والمعادلة (١٢) يمكن كتابتها على الصورة :

$$\text{ع}^{٢}_{\text{ص/س}} = \frac{\text{مجـ ص}^{٢} - \text{أ مجـ ص} - \text{ب مجـ س ص}}{\text{ن} - ٢} \qquad (١٤)$$

وسوف نقوم الآن بحساب الخطأ المعيارى للتقدير بأخذ الجذر التربيعى الموجب للمعادلة(١٤).

سبق أن حسبنا معادلة خط انحدار ص/ س لبيانات العينة ووجدناها كالآتى :

$$\hat{\text{ص}} = ٤,٩١٥ + ٢,٢١٧ \text{ س}$$

من هذه المعادلة فإن : أ = ٤,٩١٥ ، ب = ٢,٢١٧

وبالاستعانة بالقيم الواردة بجدول المثال رقم (١) نجد أن :

$$\text{ع}^{٢}_{\text{ص/س}} = \frac{٢٨٣٢ - (٤,٩١٥)(١٦٠) - (٢,٢١٧)(٩٠٢)}{١٠ - ٢}$$

$$= \frac{٢٨٣٢ - ٧٨٦,٤٠٠ - ١٩٩٩,٧٣٤}{٨}$$

$$= \frac{٤٥,٨٦٦}{٨} = ٥,٧٣٣٢٥$$

$$\text{ع}_{\text{ص/س}} = \sqrt{٥,٧٣٣٢٥} = \underline{\underline{٢,٣٩٤}}$$

---

(١) ينبغى على الدارس أن يفرق بين الخطأ المعيارى للتقدير والذى يقيس درجة تشتت القراءات ص حول خط الانحدار ، وبين الانحراف المعيارى لـ ص والذى يقيس درجة تشتت القراءات ص حول وسطه الحسابى $\bar{\text{ص}}$. والذى يعرف كالآتى :

$$\text{ع}_{\text{ص}} = \sqrt{\frac{\text{مجـ}(\text{ص} - \bar{\text{ص}})^{٢}}{\text{ن} - ١}} = \sqrt{\frac{\text{مجـ ص}^{٢} - \text{ن } \bar{\text{ص}}^{٢}}{\text{ن} - ١}}$$

التغير المفسر وغير المفسر :

Explained and Unexplained Variation

عندما يقوم الباحث بدراسة التغير (أو الاختلاف) الكلى للمتغير التابع ص فإنه يقوم بحساب مجموع مربعات الفروق بين قيم ص ووسطها الحسابى $\overline{ص}$ ويطلق على هذا المجموع بالتغير (أو الاختلاف) الكلى total variation ويعرف كالآتى :

$$مجـ (ص - \overline{ص})^2$$

هذا ويمكن تجزئة partition هذه الكمية إلى جزئين :

الجزء الأول : ويعزى إلى علاقة الانحدار بين المتغير التابع ص والمتغير المستقل س . ويمثل هذا الجزء مجموع مربعات الفروق بين قيم المتغير التابع المقدرة من معادلة خط الانحدار $\hat{ص}$ والوسط الحسابى $\overline{ص}$. ويطلق على هذا الجزء بالتغير المفسر explained variation بالعلاقة الخطية بين المتغيرين موضوع الدراسة . ويعرف كالآتى : $مجـ (\hat{ص} - \overline{ص})^2$

الجزء الثانى : ويعزى إلى عوامل أخرى ليس فى استطاعة الباحث تحديدها أو التنبؤ بها unpredictable . ويمثل هذا الجزء مجموع مربعات الفروق بين قيم المتغير التابع الفعلية أى ص وبين قيمة المقدرة من معادلة خط الانحدار أى $\hat{ص}$. ويطلق على هذا الجزء بالتغير غير المفسر (أو مجموع مربعات الخطأ) unexplained variation ويعرف كالآتى :

$$مجـ (ص - \hat{ص})^2$$

وعلى ذلك يكون لدينا العلاقة الآتية :

$$مجـ (ص - \overline{ص})^2 = مجـ (ص - \hat{ص})^2 + مجـ (\hat{ص} - \overline{ص})^2 \qquad (١٥)$$

التغير الكلى = التغير غير المفسر + التغير المفسر .

، كما سنوضح فيما يلي صحة العلاقة (١٥) على بيانات المثال (١)

التغير الكلي		التغير غير المفسر		التغير المفسر	
(ص - ص̄)	(ص - ص̄)²	(ص - ص̂)	(ص - ص̂)²	(ص̂ - ص̄)	(ص̂ - ص̄)²
٣	٩	٠,٨٧٣	٠,٦١٣	٢,٢١٧	٤,٩١٥
٥	٢٥	٢,٧٨٣	٧,٧٤٥	٢,٢١٧	٤,٩١٥
٦	٣٦	١,٥٦٦	٢,٤٥٢	٤,٤٣٤	١٩,٦٦٠
١٠-	١٠٠	١,١٣٢-	١,٢٨١	٨,٨٦٨-	٧٨,٦٤١
٢-	٤	٢,٤٣٤	٥,٩٢٤	٤,٤٣٤-	١٩,٦٦٠
٣-	٩	٣,٠٠٠-	٩,٠٠٠	-	-
٧-	٤٩	٢,٥٦٦-	٦,٥٨٤	٤,٤٣٤-	١٩,٦٦٠
٠	٠	-	-	-	-
٦	٣٦	٢,٨٦٨-	٨,٢٢٥	٨,٨٦٨	٧٨,٦٤١
٢	٤	٢,٠٠٠	٤,٠٠٠	-	-
٢٢	٢٧٢	٩,٥٦٦	٤٥,٨٢٤	١٧,٧٣٦	٢٢٦,٠٩٤
٢٢-		٩,٥٦٦-		١٧,٧٣٦-	
صفر		صفر		صفر	

التغير الكلـــــي = مجـ(ص - ص̄)² = ٢٧٢

التغير غير المفسر = مجـ(ص - ص̂)² = ٤٦ (بعد التقريب)

التغير المفـــسر = مجـ(ص̂ - ص̄)² = ٢٢٦ (بعد التقريب)

ومن الواضح أن :

مجـ(ص - ص̄)² = مجـ(ص - ص̂)² + مجـ(ص̂ - ص̄)²

٢٧٢ = ٤٦ + ٢٢٦ = ٢٧٢

الطريقة المختصرة :

من الواضح أن حساب كل من التغير الكلى والتغير المفسر والتغير غير المفسر باستخدام الصيغ التعريفية السابقة يتطلب الكثير من الوقت والجهد ، لذلك يفضل استخدام الصيغ الحسابية الآتية فى حساب كل منهما :

التغير الكلــى = مجـ (ص − ص̄)^{2} — الصيغة التعريفية

= مجـ ص^{2} − ن ص̄^{2} — الصيغة الحسابية

التغير المفسر = مجـ (ص̂ − ص̄)^{2} — الصيغة التعريفية

= ب^{2} مجـ (س − س̄)^{2}

= ب^{2} { مجـ س^{2} − ن س̄^{2} } — الصيغة الحسابية الأولى

أو = ب { مجـ (س − س̄) (ص − ص̄) }

= ب { مجـ س ص − ن س̄ ص̄ } — الصيغة الحسابية الثانية

التغير غير المفسر = التغير الكلى − التغير المفسر — أو

مجـ (ص − ص̂)^{2} = بسط تباين خط الانحدار

= مجـ ص^{2} − أ مجـ ص − ب مجـ س ص — الصيغة الحسابية

وبالتطبيق على المثال (١) :

التغير الكلــى = مجـ ص^{2} − ن ص̄^{2}

= ٢٨٣٢ − ١٠ (١٦)^{2} = ٢٨٣٢ − ٢٥٦٠ = ٢٧٢

التغير المفســر = ب^{2} مجـ (س − س̄)^{2} = ب^{2} { مجـ س^{2} − ن س̄^{2} }

= (٢,٢١٧)^{2} { ٢٩٦ − ١٠ (٥)^{2} } = ٢٢٦,٠٩٤ ≃ ٢٢٦

التغير غير المفسر = التغير الكلى − التغير المفسر

= ٢٧٢ − ٢٢٦ = ٤٦ — وهى نفس النتائج السابقة .

معامل التحديد Coefficient of Determination

يطلق على النسبة بين التغير المفسر إلى التغير الكلى بمعامل التحديد للعينة the sample coefficient of determination ويرمز له بالرمز ر٢ وهذا يعنى أن :

$$ر^2 = \frac{\text{التغير المفسر}}{\text{التغير الكلى}} = \frac{\text{مجـ}(\hat{ص} - \bar{ص})^2}{\text{مجـ}(ص - \bar{ص})^2} \qquad (١٦)$$

ومعامل التحديد للعينة – كتقدير لمعامل التحديد للمجتمع – يقيس مدى جودة توفيق goodness - of - fit خط الانحدار لبيانات العينة . ومن المعلوم أنه كلما كبر حجم العينة كلما كانت ر٢ تقديراً متسقاً وغير متحيزاً لمعامل التحديد فى المجتمع .

والجذر التربيعى لمعامل التحديد ر٢ عبارة عن معامل الارتباط للعينة وتكون إشارة معامل الارتباط ر فى هذه الحالة هى نفس إشارة معامل الانحدار ب فى معادلة خط انحدار ص/ س وهذا يعنى أن :

$$ر = \pm \sqrt{ر^2}$$

وقيمة معامل التحديد تتراوح بين الصفر والواحد الصحيح حيث يكون :

أولاً : معامل التحديد = ١ :

إذا كان التغير غير المفسر مساوياً للصفر ، فإن التغير الكلى سوف يكون مساوياً للتغير المفسر ، وعلى ذلك يكون معامل التحديد مساوياً للواحد الصحيح .

وهذا يعنى أن جميع نقط الشكل الانتشارى تقع تماماً على الخط المستقيم . وفى هذه الحالة يكون التوفيق تاماً .

ثانياً : – معامل التحديد = صفراً

إذا كان التغير المفسر مساوياً للصفر ، فإن خط الانحدار فى هذه الحالة سوف يكون أفقياً ومارأ بالوسط الحسابى $\bar{ص}$. وهذا يعنى أن مجموع التغير يكون جميعه غير مفسراً ، وعلى ذلك يكون معامل التحديد مساوياً للصفر . وهذا هو الحد الأدنى لقيمة معامل التحديد لأن هذا المعامل لا يكون أبداً أقل من الصفر (أى أنه لا يكون أبداً مساوياً لقيمة سالبة) .

وفى الحالات الأخرى عندما يكون معامل التحديد قريباً من الواحد الصحيح فإن نقط الشكل الانتشارى سوف تكون مركزة بالقرب من خط الانحدار وتأخذ مظهر الخط المستقيم بما يعنى وجودة علاقة قوية بين المتغيرين س ، ص .

وعندما يكون معامل التحديد قريباً من الصفر فإن نقط الشكل الانتشارى سوف تكون مبعثرة فى جميع أنحاء الشكل البيانى بشكل غير منتظم بما يعنى ضعف أو عدم وجود علاقة بين المتغيرين س ، ص موضوع الدراسة .

وبالنسبة للمثال (١) فإن :

$$\text{معامل التحديد } ر^{٢} = \frac{\text{التغير المفسر}}{\text{التغير الكلى}} = \frac{٢٢٦}{٢٧٢} = ٠,٨٣١ .$$

وهذا يشير إلى أن ٨٣ ٪ من التغيرات فى المتغير التابع ص أمكن تفسيرها بواسطة التغيرات فى المتغير المستقل س والباقى وقدره ١٧ ٪ من التغيرات فى ص لا تفسرها العلاقة الخطية بين المتغيرين .

$$\text{معامل الارتباط : } ر = \pm \sqrt{ر^{٢}} = \pm \sqrt{٠,٨٣١} = ٠,٩١٢ .$$

(اشارة ب فى معادلة خط الانحدار ص/ س موجبة)

صورة أخرى لحساب معامل التحديد :

من المعادلة (١٦) ذكرنا أن :

$$ر^{٢} = \frac{\text{التغير المفسر}}{\text{التغير الكلي}} = \frac{مجـ(\hat{ص} - \bar{ص})^{٢}}{مجـ(ص - \bar{ص})^{٢}}$$

$$= \frac{مجـ(ص - \bar{ص})^{٢} - مجـ(ص - \hat{ص})^{٢}}{مجـ(ص - \bar{ص})^{٢}}$$

$$= \frac{مجـ(ص - \bar{ص})^{٢}}{مجـ(ص - \bar{ص})^{٢}} - \frac{مجـ(ص - \hat{ص})^{٢}}{مجـ(ص - \bar{ص})^{٢}}$$

$$= ١ - \frac{مجـ(ص - \hat{ص})^{٢}}{مجـ(ص - \bar{ص})^{٢}} = ١ - \frac{\text{التغير غير المفسر}}{\text{التغير الكلي}} \quad (١٧)$$

وبالنسبة للمثال (١) فإن :

$$\text{معامل التحديد ر} = ١ - \frac{٤٩}{٢٧٢} = ١ - ٠,١٦٩ = ٠,٨٣١$$

العلاقة بين معامل الارتباط ر ومعامل الانحدار ب لخط انحدار ص / س

من دراستنا للارتباط نعلم أن معامل الارتباط يحسب من المعادلة :

$$ر = \frac{مجـ(س - \bar{س})(ص - \bar{ص})}{\sqrt{مجـ(س - \bar{س})^{٢}} \sqrt{مجـ(ص - \bar{ص})^{٢}}}$$

ومن دراستنا للانحدار نعلم أن معامل انحدار ص/ س يحسب من المعادلة :

$$ب = \frac{مجـ(س-\bar{س})(ص-\bar{ص})}{مجـ(س-\bar{س})^2}$$

وهذا يعنى أن بسط معامل الارتباط هو نفسه بسط معامل الانحدار وسوف نوضح فيما يلى العلاقة بين ب ، ر

$$مجـ(س-\bar{س})^2 = \sqrt{مجـ(س-\bar{س})^2}\sqrt{مجـ(س-\bar{س})^2}$$

$$\therefore ب = \frac{مجـ(س-\bar{س})(ص-\bar{ص})}{\sqrt{مجـ(س-\bar{س})^2}\sqrt{مجـ(ص-\bar{ص})^2}}$$

بضرب المعادلة السابقة بسطاً ومقاماً فى الكمية $\sqrt{مجـ(ص-\bar{ص})^2}$ ينتج أن :

$$\therefore ب = \frac{مجـ(س-\bar{س})(ص-\bar{ص})}{\sqrt{مجـ(س-\bar{س})^2}\sqrt{مجـ(ص-\bar{ص})^2}} \times \frac{\sqrt{مجـ(ص-\bar{ص})^2}}{\sqrt{مجـ(س-\bar{س})^2}}$$

$$\therefore ب = ر \times \frac{ع_{ص}}{ع_{س}} \qquad (18)$$

والمعادلة السابقة تمكننا من حساب معامل الانحدار ب بمعلومية معامل الارتباط ر والعكس.

<u>وسنوضح العلاقة (١٨) بالنسبة لبيانات المثال (١) :</u>

سبق أن حسبنا قيمة معامل الارتباط ووجدنا أن : ر = ٠,٩١٢ .

، كما حسبنا قيمة معامل انحدار ص/ س ووجدنا أن : ب = ٢,٢١٧

سنقوم الآن بحساب كل من $ع_{ص}$ ، $ع_{س}$ كالآتى :

$$ع_ص = \sqrt{\frac{مجـ(ص - \bar{ص})^٢}{ن - ١}} = \sqrt{\frac{٢٧٢}{٩}} = \sqrt{٣٠,٢٢٢٢} = ٥,٥٠$$

$$ع_س = \sqrt{\frac{مجـ(س - \bar{س})^٢}{ن - ١}} = \sqrt{\frac{٤٦}{٩}} = \sqrt{٥,١١١١} = ٢,٢٦$$

وبالتعويض بالقيم السابقة فى المعادلة : $ب = ر \times \frac{ع_ص}{ع_س}$ ينتج أن :

$$ب = ٠,٩١١ \times \frac{٥,٥}{٢,٦} = ٠,٩١١ \times ٢,٤٣٣٦ = \underline{\underline{٢,٢١٧}}$$

وهى نفس النتيجة السابقة .

إيجاد معامل خط انحدار ص/ س بدلالة الارتباط :

من المعادلتين (٢) ، (٦) فإنه يمكن كتابة خط انحدار ص/ س الآتية :

كما يلى : $\hat{ص} = أ + ب س$

$$\hat{ص} - \bar{ص} = ب(س - \bar{س}) \qquad (١٩)$$

وبالتعويض بما تساويه ب فى المعادلة (١٨) فى المعادلة السابقة ، نجد أن :

$$\hat{ص} - \bar{ص} = ر \times \frac{ع_ص}{ع_س}(س - \bar{س}) \qquad (٢٠)$$

$$\text{أو} \quad \hat{ص} = ر \times \frac{ع_ص}{ع_س}(س - \bar{س}) + \bar{ص} \qquad (٢١)$$

ومن الواضح أن المعادلة (٢١) تمكننا من تقدير قيمة ص بمعلومية معامل الارتباط .

سنوضح فيما يلى العلاقة (٢٠) بالنسبة لبيانات المثال (١) على النحو الآتى :

$$\hat{ص} = ٠,٩١٢ \times \frac{٥,٥٠}{٢,٢٦}(س - ٥) + ١٦$$

$$= ٢,٢١٧(س - ٥) + ١٦ = ٢,٢١٧ س - ١١,٠٨٥ + ١٦$$

$$\therefore \hat{ص} = ٤,٩١٥ + ٢,٢١٧ س$$

وهى نفس معادلة خط الانحدار ص/ س السابق حسابها .

مثال تعام

باستخدام بيانات الجدول الآتى الخاصة بالمتغيرين س ، ص :

س	٢٤	١٨	١٧	٢٠	١٩	٢٢	٢١	١٩
ص	٢٧	٣٥	٣٧	٣٤	٣٣	٢٦	٢٩	٢٥

أولاً : ارسم شكل الانتشار للبيانات السابقة ، وحدد بالفحص النظرى عما إذا كانت توجد علاقة خطية تقريبية بين المتغيرين س ، ص .

ثانياً : أوجد خط انحدار المربعات الصغرى للبيانات الخاصة بالمتغيرين س ، ص . ثم فسر معنى ثوابت الانحدار .

ثالثاً : ارسم خط الانحدار على التمثيل البيانى لنقط شكل الانتشار ، والوارد فى البند أولاً ، ثم وضح ما إذا كانت معادلة خط الانحدار تصف بكفاية العلاقة الخطية بين المتغيرين س ، ص .

رابعاً : احسب التغير الكلى ، والتغير المفسر ، والتغير غير المفسر ، ثم استخدم هذه البيانات فى حساب :(١) التباين والخطأ المعيارى للتقدير .

(٢) معامل التحديد ومعامل الارتباط الخطى البسيط .

خامساً : قدر قيمة ص عندما تكون س = ١٨ ، س = ٢٧ ، س = س̄ .

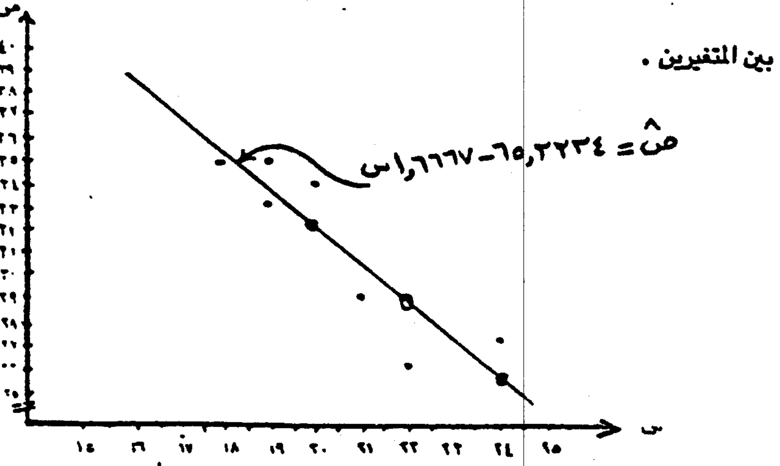
سادساً : احسب معامل الارتباط الخطى البسيط باستخدام صيغة معامل بيرسون للارتباط للبيانات الخاصة بالمتغيرين س ، ص . ثم قارن بين النتيجة التى حصلت عليها بالقيمة الواردة فى البند رابعاً .

سابعاً : أحسب معامل الارتباط الخطى البسيط بمعلومية معامل الانحدار المحتسب فى البند ثانياً .



الحــــــل

أولاً : من الفحص النظرى لنقط شكل الانتشار يتبين وجود علاقة خطية عكسية تقريبية بين المتغيرين .



ثانياً : لحساب معادلة خط انحـــدار ص/ س ننشئ الجدول الآتى :

س	س	س ص	س٢	ص٢
٢٤	٢٧	٦٤٨	٥٧٦	٧٢٩
١٨	٣٥	٦٣٠	٣٢٤	١٢٢٥
١٧	٣٧	٦٢٩	٢٨٩	١٣٦٩
٢٠	٣٤	٦٨٠	٤٠٠	١١٥٦
١٩	٣٣	٦٢٧	٣٦١	١٠٨٩
٢٢	٢٦	٥٧٢	٤٨٤	٦٧٦
٢١	٢٩	٦٠٩	٤٤١	٨٤١
١٩	٣٥	٦٦٥	٣٦١	١٢٢٥
١٦٠	٢٥٦	٥٠٦٠	٣٢٣٦	٨٣١٠

$\bar{س} = \frac{١٦٠}{٨} = ٢٠$ $\bar{ص} = \frac{٢٥٦}{٨} = ٣٢$

معادلة انحدار ص/ س تأخذ الصورة الاتية : ص = أ + ب س حيث :

$$ب = \frac{مجـ س ص - ن \bar{س} \bar{ص}}{مجـ س^٢ - ن \bar{س}^٢} = \frac{٥٠٦٠ - ٨(٢٠)(٣٢)}{٣٢٣٦ - ٨(٢٠)^٢}$$

$$= \frac{٥٠٦٠ - ٥١٢٠}{٣٢٣٦ - ٣٢٠٠} = \frac{-٦٠}{٣٦} = - ١,٦٦٦٦٧$$

$$أ = \bar{ص} - ب \bar{س}$$

$$= ٣٢ - (-١,٦٦٦٦٧)(٢٠) = ٣٢ + ٣٣,٣٣٣٤ = ٦٥,٣٣٣٤$$

وتكون معادلة خط الانحدار المقدرة (ص/ س) مقربة إلى ٤ أرقام عشرية كالآتي :

$$\hat{ص} = ٦٥,٣٣٣٤ - ١,٦٦٦٧ س$$

ويمكن تفسير ثوابت الانحدار كالآتي :

أ = ٦٥,٣٣٣٤ تعبر عن القيمة المتوقعة للمتغير التابع ص عندما س = صفر .

، ب = - ١,٦٦٦٧ تعنى أنه عندما يزداد المتغير المستقل س بوحدة واحدة فإن المتغير التابع ص يقل فى المتوسط بمقدار ١,٦٦٦٧ وحدة . وهذا يعنى أن العلاقة بين المتغيرين س ، ص علاقة عكسية .

ثالثاً : لرسم خط الانحدار المقدر ، فإنه يلزمنا تعيين نقطتين ، إلا أنه من المفضل أخذ نقطة ثالثة للتحقق من دقة النقطتين الأخرتين .

فإذا أعطينا س قيماً عددية ولتكن ٢٠ ، ٢٢ ، ٢٤ ، فإننا نعوض بها فى معادلة الانحدار للحصول على قيم ص المناظرة لها ، ثم ندون النتائج فى الجدول :

س	٢٠	٢٢	٢٤
ص	٣٢	٢٨,٧	٢٥,٣

والخطوة التالية هى رصد هذه النقط الثلاثة على شكل الانتشار ثم توصل فيما بينها نحصل على خط الانحدار الممثل للعلاقة بين المتغيرين س ، ص كما فى الشكل البيانى .

من الفحص النظرى ، يتضح أن النقط موزعة حول خط الانحدار ومتقاربة منه ، مما يعنى وجود علاقة خطية عكسية بين المتغيرين س ، ص .

رابعاً : سوف نقوم بحساب التغير الكلى والتغير المفسر والتغير غير المفسر بالطريقة المختصرة .

التغير الكلـــى = مجـ (ص − $\overline{ص}$)^{٢} = مجـ ص^{٢} − ن $\overline{ص}^{٢}$

= ٨٣١٠ − ٨ (٣٢)^{٢} = ٨٣١٠ − ٨١٩٢ = ١١٨

التغير المفسر = مجـ (ص − $\overline{ص}$)^{٢} = ب^{٢} مجـ (س − $\overline{س}$)^{٢} = ب^{٢} [مجـ س^{٢} − ن $\overline{س}^{٢}$]

=(−١,٦٦٦٧)^{٢} [٣٢٣٦ − ٨(٢٠)^{٢}] ≃ ١٠٠

التغير غير المفسر = التغير الكلى − التغير المفسر = ١٨٨ − ١٠٠ = ١٨

ومن النتائج السابقة يمكن حساب المطلوب :

(١) التباين $ع^{٢}_{ص/س}$ = $\frac{مجـ (ص − \hat{ص})^{٢}}{ن − ٢}$ = $\frac{مجموع\ مربعات\ الخطأ}{ن − ٢}$

= $\frac{١٨}{٨ − ٢}$ = $\frac{١٨}{٦}$ = ٢

الخطأ المعيارى للتقدير = $\sqrt{٢}$ = ١,٧٣٢١

(٢) معامل التحديد = $\frac{مجـ (\hat{ص} − \overline{ص})^{٢}}{مجـ (ص − \overline{ص})^{٢}}$ = $\frac{التغير\ المفسر}{التغير\ الكلى}$ = $\frac{١٠٠}{١١٨}$ = ٠,٨٤٧٥

معامل الارتباط = ± $\sqrt{٠,٨٤٧٥}$ = − ٠,٩٢٠٦

(الاشارة هنا سالبة لأن اشارة ب فى معادلة خط الانحدار سالبة)

خامساً :

عندما س = ١٨ فإن $\hat{ص}$ = ٦٥,٣٣٣٤ − ١,٦٦٦٧ (١٨) = ٣٥,٣٣٢٨

. عندما س = ٢٧ فإن $\hat{ص}$ = ٦٥,٣٣٣٤ − ١,٦٦٦٧ (٢٧) = ٢٠,٣٣٢٥

. عندما س = $\overline{س}$ = $\frac{مجـ س}{ن}$ = $\frac{١٦٠}{٨}$ = ٢٠

فإن $\hat{ص}$ = ٦٥,٣٣٣٤ − ١,٦٦٦٧ (٢٠) = ٣١,٩٩٩٤

سابعاً : لحساب معامل بيرسون للارتباط الخطي البسيط باستخدام بيانات المتغير س ، ص نستخدم المجاميع والمربعات وحواصل الضرب من الجدول الوارد في البند ثانياً .

$$\text{ر} = \frac{\text{مجـ س ص} - \text{ن}\ \overline{\text{س}}\ \overline{\text{ص}}}{\sqrt{\text{مجـ س}^{٢} - \text{ن}\ \overline{\text{س}}^{٢}}\ \sqrt{\text{مجـ ص}^{٢} - \text{ن}\ \overline{\text{ص}}^{٢}}}$$

$$= \frac{\text{٥٠٦٠} - \text{٨}(\text{٢٠})(\text{٣٢})}{\sqrt{\text{٣٢٣٦} - \text{٨}(\text{٢٠})^{٢}}\ \sqrt{\text{٨٣١٠} - \text{٨}(\text{٣٢})^{٢}}} = \frac{\text{٥٠٦٠} - \text{٥١٢٠}}{\sqrt{\text{٣٦}}\ \sqrt{\text{١١٨}}}$$

$$= \frac{-\text{٦٠}}{\sqrt{\text{٣٦} \times \text{١١٨}}} = -\frac{\text{٦٠}}{\text{٦٥,١٧٦٦٨}} = -\text{٠,٩٢٠٦}$$ (كالسابق)

سابعاً : إيجاد الارتباط بمعلومية الانحدار :

من البند ثانياً نجد أن معامل الانحدار ب = − ١,٦٦٦٧

ولحساب الارتباط بمعلومية هذه القيمة نستخدم المعادلة الآتية :

$$\text{ر} = \text{ب} \cdot \frac{\text{ع}_{\text{س}}}{\text{ع}_{\text{ص}}} = (-\text{١,٦٦٦٧}) \frac{\text{ع}_{\text{س}}}{\text{ع}_{\text{ص}}}$$ حيث :

$$\text{ع}_{\text{س}} = \sqrt{\frac{\text{مجـ}(\text{س} - \overline{\text{س}})^{٢}}{\text{ن} - \text{١}}} = \sqrt{\frac{\text{مجـ س}^{٢} - \text{ن}\ \overline{\text{س}}^{٢}}{\text{ن} - \text{١}}}$$

$$= \sqrt{\frac{\text{٣٢٣٦} - \text{٨}(\text{٢٠})^{٢}}{\text{٨} - \text{١}}} = \sqrt{\frac{\text{٣٦}}{\text{٧}}} = \text{٢,٢٦٧٨}$$

$$\text{ع}_{\text{ص}} = \sqrt{\frac{\text{مجـ}(\text{ص} - \overline{\text{ص}})^{٢}}{\text{ن} - \text{١}}} = \sqrt{\frac{\text{مجـ ص}^{٢} - \text{ن}\ \overline{\text{ص}}^{٢}}{\text{ن} - \text{١}}}$$

$$= \sqrt{\frac{\text{٨٣١٠} - \text{٨}(\text{٣٢})^{٢}}{\text{٨} - \text{١}}} = \sqrt{\frac{\text{١١٨}}{\text{٧}}} = \text{٤,١٠٥٧}$$

$$\therefore \text{ر} = (-\text{١,٦٦٦٧}) \left[\frac{\text{٢,٢٦٧٨}}{\text{٤,١٠٥٧}}\right] = -\text{٠,٩٢٠٦}$$ (كالسابق)

تطبيقـــــات الباب الخامس

١ - من البيانات الآتية الخاصة بالمتغيرين س ، ص :

س	٢	٤	٦	٨	١٠
ص	٧	٧	٦	٤	١

(أ) ارسم شكل الانتشار الذى يمثل العلاقة بين المتغيرين س ، ص

(ب) احسب معادلة خط انحدار ص/ س .

(جـ) احسب الخطأ المعيارى لخط انحدار ص/ س .

(د) احسب التغير المفسر والتغير غير المفسر فى ص بواسطة خط انحدار ص/ س .

(هـ) احسب معامل التحديد ومعامل الارتباط بين المتغيرين س ، ص .

٢ - الجدول الآتى يوضح الدخل الشهرى (س) والاستهلاك (ص) لسبعة أسر اختيرت بطريقة عشوائية :

س	١٩	٢٦	٢٩	٣٠	٣٢	٣٥	٣٩
ص	١٨	١٩	٢٩	٢٧	٢٠	٣٠	٣٢

والمطلوب : (أ) رسم شكل الانتشار الذى يوضح العلاقة بين المتغيرين س ، ص

(ب) حساب معادلة خط انحدار ص/ س .

(جـ) حساب التباين والخطأ المعيارى للتقدير .

(د)تقدير قيم ص عندما س = ٢٨ ، س = ٣٧ ، س = $\overline{\text{س}}$.

(هـ) حساب معامل التحديد ومعامل الارتباط بين المتغيرين س ، ص .

٣ - الجدول الآتى يوضح الخبرة بالسنوات (س) وحجم المبيعات بعشرات الآلاف من الجنيهات (ص) لمجموعة مكونة من عشرة بائعين :

س	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠
ص	٤	٣	٦	٥	٧	٩	١٠	١٠	١٤	١٢

والمطلوب : (أ) حساب معادلة خط انحدار ص/ س بفرض أن العلاقة بين المتغيرين علاقة خطية .

(ب) حساب التباين والخطأ المعيارى لخط انحدار ص/ س .

(جـ) معامل التحديد ومعامل الارتباط بين المتغيرين س ، ص .

٤ - الجدول الآتى يوضح الطول بالبوصة (س) والوزن بالرطل (ص) لعشرة طلاب من كلية معينة :

س	٦٤	٧١	٧٠	٦٨	٧٠	٦٨	٧٣	٦٩	٧٠	٦٧
ص	١٥٦	١٧٨	١٧٥	١٦٧	١٦٤	١٦٠	١٩٨	١٦٩	١٨٣	١٨٠

والمطلوب : (أ) حساب معادلة خط انحدار ص/ س بفرض أن العلاقة بين المتغيرين علاقة خطية .

(ب) حساب التباين والخطأ المعيارى لخط انحدار ص/ س .

(جـ) تقدير قيم ص عندما س = ٦٥ ، س = ٧٢ ، س = $\overline{س}$

(جـ) معامل التحديد ومعامل الارتباط بين المتغيرين س ، ص .

الباب السادس

الأرقام القياسية

الأرقام القياسية Index Numbers

الفصل الأول

تمهيد

تعتبر الأرقام القياسية وسيلة من الوسائل الإحصائية الهامة لقياس التغيرات النسبية التى تحدث بالنسبة إلى الظواهر الاقتصادية أو التجارية أو الإدارية أو الاجتماعية : كتغيرات الأسعار والإنتاج والتجارة الداخلية والتجارة الخارجية والأجور والعمالة والبطالة وعدد السكان ... إلخ ، بالنسبة إلى فترتين زمنيتين مختلفتين أو مكانين مختلفين .

كما تظهر أهمية الأرقام القياسية فى أن بعض الطرق المتعلقة بعمل تنبؤات أو اتخاذ قرارات مستقبلية ، كثيراً ما تطبق في صورة أرقام قياسية . ففى تحليل السلاسل الزمنية – على سبيل المثال – فقد يكون المتغير التابع أو المتغير المستقل أو كليهما فى صيغة أرقام قياسية . كما أنه كثيراً ما نجد أن المؤسسات التجارية أو الهيئات أو الحكومات لا ترغب فى نشر البيانات المتصلة بظروفها الاقتصادية أو المالية ولكنها لا تتردد فى إعطاء هذه البيانات فى صيغة نسب مئوية .

وهذه الوظائف وغيرها من الوظائف الهامة للأرقام القياسية – والتى سنتناولها بالدراسة فيما بعد – جعلت الأرقام القياسية من الأدوات الإحصائية الهامة فى التحليل الإحصائى .

تعريف :

الرقم القياسى فى أبسط صورة عبارة عن نسبة مئوية ، تستخدم لقياس التغير الذى يطرأ على ظاهرة ما بالنسبة إلى فترتين زمنيتين مختلفتين أو مكانين مختلفتين .

ومن هذا التعريف ، يتضح أن الوظيفة الأساسية للرقم القياسى البسيط هو تحويل الكميات المطلقة للظاهرة موضوع الدراسة إلى أرقام نسبية relative numbers حتى يمكن مقارنة التغيرات التي تحدث للظاهرة بالنسبة إلى عدة أزمنة مختلفة أو بالنسبة الي عدة مناطق مختلفة

وفى دراستنا لموضوع الأرقام القياسية ، سوف نركز اهتمامنا بالنسبة للأرقام القياسية الخاصة بكل من الأسعار والكميات هذا ، ويطلق على الرقم القياسي البسيط الذى يقيس التغير بالنسبة لسعر (أو كمية) سلعة واحدة بالمنسوب . ويترتب على ذلك أن يكون :

$$\text{منسوب السعر} = \frac{\text{سعر السلعة فى سنة المقارنة}}{\text{سعر السلعة فى سنة الأساس}} \times ١٠٠ \qquad (١)$$

$$\text{منسوب الكمية} = \frac{\text{كمية السلعة فى سنة المقارنة}}{\text{كمية السلعة فى سنة الأساس}} \times ١٠٠ \qquad (٢)$$

هذا ، ولما كانت قيمة أى سلعة عبارة عن حاصل ضرب سعر هذه السلعة فى كميتها . . على ذلك فإن :

$$\text{منسوب القيمة} = \frac{\text{قيمة السلعة فى سنة المقارنة}}{\text{قيمة السلعة فى سنة الأساس}} \times ١٠٠ \qquad (٣)$$

الرموز المستخدمة .

لتجنب التكرار فى تعريف معنى الرموز المستخدمة ، سوف نعرف فيما يلى الرموز التى سوف نستخدمها فى الوقت الحالى :

- $ع_٠$: سعر السلعة فى سنة (أو فترة) الأساس .
- $ع_١$: " " " " (أوفترة) المقارنة .
- $ك_٠$: كمية " " " (أوفترة) الأساس .
- $ك_١$: " " " " (أوفترة) المقارنة .
- س : منسوب السعر (بالنسبة لسلعة واحدة) .
- ك : منسوب الكمية (بالنسبة لسلعة واحدة) .

أنواع الأرقام القياسية :

جدير بالذكر أن هناك كثيراً من الاعتبارات النظرية والتطبيقية التى تتعلق بتركيب وعمل الأرقام القياسية . وقبل مناقشة هذه الاعتبارات - والتي سنتناولها بالشرح والتحليل فيما بعد - فإننا سوف نركز الآن على النواحي الحسابية والفنية الخاصة بحساب الأرقام القياسية

وغني عن البيان أن الأرقام القياسية لها أنواع متعددة ومن الممكن أن تكون بطرق كثيرة الأمر الذي يتعذر معه إمكان تغطية جميع الأنواع المختلفة منها في هذا الكتاب وإزاء ذلك سوف نركز على أهم الأنواع الشائعة في الاقتصاد وإدارة الأعمال

وبصفة عامة ، فإن الأرقام القياسية يمكن تصنيفها إلى نوعين رئيسيين هما :

النوع الأول الأرقام القياسية البسيطة (غير المرجحة)

النوع الثاني . الأرقام القياسية المرجحة .

المبحث الأول

الأرقام القياسية البسيطة

(أ) بالنسبة للأسعار .

الرقم التجميعي البسيط للأسعار :

لحساب هذا الرقم نتبع الخطوات الآتية

١- تجمع الأسعار المختلفة للسلع المعطاة عن السنوات المختلفة لنحصل بذلك على مجـ ع ، مجـ ع١

٢- نقسم مجموع الأسعار بالنسبة إلى كل سنة على مجموع أسعار سنة الأساس .

٣ - نضرب خارج القسمة في ١٠٠ . فيكون الناتج هو الرقم التجميعي البسيط للأسعار

وهذا يعني أن

$$\text{الرقم التجميعي البسيط للأسعار} = \frac{\text{مجموع أسعار سنة المقارنة}}{\text{مجموع أسعار سنة الأساس}} \times ١٠٠$$

$$= \frac{\text{مجـ ع}_١}{\text{مجـ ع}} \times ١ \qquad (٤)$$

مثال (١)

الجدول الآتي يوضح أسعار أربعة سلع مختلفة أ ب حـ ، د بالجنيهات عن السنوات ١٩٨٥ ، ١٩٩٠ ، ١٩٩٥ وكذلك الكميات المستهلكة منها بآلاف الوحدات

جدول (١) أسعار وكميات بعض السلع

السلعة	وحدة القياس	الأسعار (بالجنيه)			الكميات (بالآلاف)		
		١٩٨٥	١٩٩٠	١٩٩٥	١٩٨٥	١٩٩٠	١٩٩٥
أ	الطن	٥٠	٧٠	١٧٠	٤٠	٤٠	٢٠
ب	الكيلو	٣٠	٤٢	٨٠	٦٠٠	٧٤٠	٩٦٠
جـ	الجالون	١٠	١٤	٣٦	٢٠٠	٤٥٠	٦٠٠
د	الدستة	١٨	٢٤	٤٨	١٨٠	٢٢٠	٤٨٠

والمطلوب حساب الرقم التجميعي البسيط[(١)] للأسعار لكل من سنة ١٩٩٠ و سنة ١٩٩٥ بالنسبة إلى سنة ١٩٨٥ كأساس ، وأيضاً الرقم التجميعي للأسعار لسنة ١٩٩٥ بالنسبة لسنة ١٩٩٠ كأساس .

الحـــل :

من المعادلة رقم (٤) نعلم أن :

$$\text{الرقم التجميعي البسيط للأسعار} = \frac{\text{مجموع أسعار سنة المقارنة}}{\text{مجموع أسعار سنة الأساس}} \times ١٠٠ .$$

الرقم التجميعي البسيط للأسعار سنة ١٩٩٠ بالنسبة إلى أسعار سنة ١٩٨٥ كأساس

$$= \frac{٧٠ + ٤٢ + ١٤ + ٢٤}{٥٠ + ٣٠ + ١٠ + ١٨} \times ١٠٠$$

$$= \frac{١٥٠}{١٠٨} \times ١٠٠ = ١٣٨.٩ \%$$

ونستنتج من ذلك أن أسعار السلع في سنة ١٩٩٠ قد ارتفعت في المتوسط بنسبة ٣٨.٩ ٪ عما كانت علية أسعارها في سنة ١٩٨٥

(١) اذا كانت وحدات النقود المعبرة عن أسعار السلعة مختلفة باختلاف وحدات هذه السلع بمعنى أنه ، اذا ذكر أن سعر الوحدة من السلعة أ بالجنيهات ، وأن سعر الوحدة من السلعة ب قد ذكر بالقروش وهكذا ، فإن الامر يستلزم قبل حساب الرقم التجميعي البسيط للأسعار توحيد وحدات النقود لمختلف السلع الداخلة في تركيب هذا الرقم

- الرقم التجميعي البسيط لأسعار سنة ١٩٩٥ بالنسبة إلى اسعار سنة ١٩٨٥ كأساس

$$= \frac{١٧٠ + ٨٠ + ٢٦ + ٤٨}{٥٠ + ٣٠ + ١٠ + ١٨} \times ١٠٠$$

$$= \frac{٣٢٤}{١٠٨} \times ١٠٠ = ٣٠٠\%$$

- الرقم القياسي لأسعار سنة ١٩٩٥ بالنسبة لأسعار سنة ١٩٩٠ كأساس

$$= \frac{١٧٠ + ٨٠ + ٢٦ + ٤٨}{٧٠ + ٤٢ + ١٤ + ٢٤} \times ١٠٠$$

$$= \frac{٣٢٤}{١٥٠} \times ١٠٠ = ٢١٦\%$$

ملاحظات على الرقم التجميعي البسيط للأسعار :

أوجدنا الرقم التجميعي البسيط للأسعار دون أن ندخل في اعتبارنا الأهمية النسبية للسلع المختلفة بالنسبة إلى بعضها البعض . فمن الجائز أن تكون السلعه ب أهم بكثير من بقية السلع الأخرى ، ومع ذلك فقد صرفنا النظر عن هذه الأهمية وعاملنا جميع السلع الداخلة في تركيب الرقم القياسي على قدم المساواة . فإنه ليس من المعقول - على سبيل المثال - أن نساوي بين الأهمية النسبية للقطن أو الصلب في الاقتصاد القومي مع الأهمية النسبية للمربى أو الترمس ونضعها جميعاً على قدم المساواة . وهذا يمثل العيب الأساسي لهذا الرقم

ويعاب أيضاً على الرقم التجميعي البسيط للأسعار ، تأثره بوحدات القياس المحتسب على أساسها أسعار السلع المختلفة . وبعبارة أخرى ، إذا أحللنا وحدة قياس محل الأخرى : بمعنى أنه إذا استخدمنا سعر الطن بدلاً من سعر الكيلو جرام ، أو سعر الأردب بدلا من سعر الكيله ، أو سعر الجالون بدلا من سعر اللتر ، أو الأجر الشهري للعامل بدلا من الأجر اليومي ، فإن ذلك سوف يؤثر على قيمة الرقم القياسي المحسوب

(ب) بالنسبة للكميات

الرقم التجميعي البسيط للكميات :

من السابق أن المعادلة الخاصة بحساب هذا الرقم هي أن

$$\text{الرقم التجميعي البسيط للكميات} = \frac{\text{مجموع كميات سنة المقارنة}}{\text{مجموع كميات سنة الاساس}} \times ١٠٠$$

$$= \frac{\text{مجـ ك}_1}{\text{مجـ ك}_\bullet} \times ١٠٠ \qquad (٥)$$

من الواضح أن هذه المعادلة لا يمكن استخدامها في حالة اختلاف وحدات قياس الكميات حيث لا يمكن جمع أطنان ، وكيلو جرامات ، وقناطير .....إلخ .

ونتيجة لذلك ، فإنه يتعذر حساب الرقم التجميعي لكميات السلع الوارده بالجدول الخاص بالمثال رقم (٦) .

وفي الحياة العملية فإنه نادراً ما تستخدم هذه المعادلة في حالة قياس التغيرات التي تحدث في الكميات .

المبحث الثانى

الأرقام التجميعية المرجحة
Weighted Aggregative Indices

أولاً - بالنسبة للأسعار

عند حسابنا للرقم التجميعى البسيط للأسعار ، صرفنا النظر عن الأهمية النسبية للسلع الداخلة فى تركيب الرقم القياسى بالنسبة إلى بعضها البعض : وعاملناها جميعاً على قدم المساواة ، الأمر الذى يؤدى إلى سيطرة السلع ذات السعر المرتفع - سيما وإن كانت قليلة الأهمية - على باقى السلع ذات السعر المنخفض وهذا يمثل عيباً من العيوب الأساسية التى توجه إلى هذا الرقم .

وللتغلب على هذا العيب الخطير ، فإن الأمر يستلزم ترجيح السلع بحسب درجة أهميتها . فتعطى السلع ذات الأهمية الكبيرة وزناً كبيراً ، أما السلع ذات الأهمية القليلة فتعطى وزناً صغيرا .

والأوزان التى تستخدم لترجيح الأرقام القياسية للأسعار هى عادة :
الكميات المستهلكة Consumed أو المنتجة Produced أو المشتراة Purchased أو المباعة Sold أو المستخدمة Used من كل سلعة ، والتى يرمز لها بالرمز ك .

ومن البديهى أن حاصل ضرب الأسعار فى الكميات (ع × ك) يمثل القيمة الكلية وهذه القيمة الكلية لا تتأثر بالوحدات المختارة ، لأن ع عبارة عن سعر الوحدة ، ك : عبارة عن عدد الوحدات .

هذا ، والأوزان الشائعة الاستخدام في ترجيح الأرقام القياسية قد تكون :

أولاً : كميات السلع فى سنة (أو فترة) الأساس ($ك_٠$) .

ثانياً : كميات السلع فى سنة (أو فترة) المقارنة ($ك_١$) .

ثالثاً : الوسط الحسابى لكميات السلع فى سنتى الأساس والمقارنة $\left(\frac{ك_٠ + ك_١}{٢} \right)$

رابعاً : الوسط الهندسى لكميات السلع فى سنتى الأساس والمقارنة $(\sqrt{ك_. ك_١})$

خامساً: الوسط التوافقى لكميات السلع فى سنتى الأساس والمقارنة

$$\left(\frac{٢}{\frac{١}{ك_.} + \frac{١}{ك_١}}\right)$$

وبالاضافة الى نظم الأوزان السابقة ، فقد يتخذ كأوزان متوسط كميات السنوات الواقعة بين سنتى الأساس والمقارنة أو كميات افتراضية hypothetical quantities

وتجدر الإشارة إلى أن كل نظام من هذه النظم من الأوزان له فوائده النظرية أو التطبيقية ، فضلا عن عيوبه . وعلى الرغم من أن تناول كل منها بالشرح والتعليق يخرج عن نطاق مناقشتنا الحالية . فإنه من الأهمية بمكان الإشارة إلى النقطتين الآتيتين :

الأولى : أن تغيير الأوزان سوف يغير بالتالى في معنى الرقم القياسى . لذلك ، فإن نظام الأوزان الذى نريد استخدامه يعتمد على السؤال الذى نبحث الإجابة عليه . وبعبارة أخرى ، يعتمد على الغرض من إنشاء الرقم القياسى المطلوب .

الثانية : من الممكن أن يعطى نوعين من الأوزان نتائج متشابهة أو متقاربة . لذلك ، فإن الأمر يستلزم أن نختار النوع الذى يتطلب جهداً يسيراً وبعداً عن العمليات الحسابية المعقدة . أو نختار النوع الذى يسمح باعطائنا تفسيرات دقيقة ومفيدة .

وإذا رمزنا للأوزان المستخدمة للسلع الداخلة في تركيب الرقم القياسى بالرموز :

$و_١ ، و_٢ ، و_٣ ، \ldots\ldots ، و_ن$

فإن الصيغة العامة للرقم التجميعى المرجح فى هذه الحالة

$$= \frac{مجـ ع_١ و}{مجـ ع_. و} \times ١٠٠$$

ونود الإشارة إلى أنه نظراً لأن النوعين الأولين من الأوزان هما الشائعا الاستخدام في الحياة العملية ، لذلك فسوف نتولى شرح كل منها بشيء من الإيجاز

وسنورد فيما يلي معادلات الأرقام التجميعية المرجحة للأسعار وفقاً لنظم الأوزان المشار إليها .

أولاً - الرقم التجميعي المرجح للأسعار باستخدام كميات سنة الأساس كأوزان :

الرقم التجميعي المرجح للأسعار باستخدام كميات سنة الأساس كأوزان :

يسمى برقم لاسبير القياسي للأسعار The Laspeyres Price Index

وإذا رمزنا لهذا الرقم بالرمز $س_{٠}$ ، فإن :

$$س_{٠} = \frac{مجـ ع_{١} ك_{٠}}{مجـ ع_{٠} ك_{٠}} \times ١٠٠ \qquad (٦)$$

وهذا يعنى أن رقم لاسبير للأسعار يمثل النسبة المئوية للقيمة الكلية لكميات سنة الأساس مقومة بأسعار سنة المقارنة إلى القيمة الكلية لنفس الكميات مقومة بأسعار سنة الأساس . وبوجه عام ، فإن رقم لاسبير يحاول الإجابة على السؤال الآتي : « ما هو التغير في القيمة الكلية لكميات سنة الأساس عند تقويمها بأسعار سنة المقارنة ؟ » وتطبيق معادلة لاسبير ، يستلزم اتباع الخطوات الثلاثة الآتية :

<u>أولاً</u> : نضرب سعر كل سلعة في سنة المقارنة في كميات هذه السلعة في سنة الأساس ، ثم نجمع حواصل الضرب المختلفة نحصل بذلك على بسط المعادلة أي على مجـ $ع_{١} ك_{٠}$.

<u>ثانياً</u> : نضرب سعر كل سلعة في سنة الأساس في كميات هذه السلعة في سنة الأساس ، ثم نجمع حواصل الضرب المختلفة نحصل بذلك عن مقام المعادلة أي على مجـ $ع_{٠} ك_{٠}$.

<u>ثالثاً</u> : نقسم القيمة الكلية لكميات سنة الأساس مقومة بأسعار سنة المقارنة والناتجة من أولا ، على القيمة الكلية لكميات سنة الأساس مقومة بأسعار سنة الأساس والناتجة من ثانياً ، ثم نضرب خارج القسمة في ١٠٠ . نحصل بذلك على رقم لاسبير المطلوب

مثـــال (٢) :

احسب رقم لاسبير القياسى للأسعار لبيانات الجدول الوارد بالمثال رقم (١) .

الحــــل :

ننشىء الجدول الآتى :

جدول (٢) جدول الأرقام التجميعية المرجحة

السلعة	الأسعار ٩٠ $ع_٠$	الأسعار ٩٥ $ع_١$	الكميات ٩٠ $ك_٠$	الكميات ٩٥ $ك_١$	$ع_٠ ك_٠$	$ع_١ ك_٠$	$ع_٠ ك_١$	$ع_١ ك_١$
أ	٧٠	١٧٠	٤٠	٣٠	٢٨٠٠	٦٨٠٠	٢١٠٠	٥١٠٠
ب	٤٢	٨٠	٧٤٠	٩٦٠	٣١٠٨٠	٥٩٢٠٠	٤٠٣٢٠	٧٦٨٠٠
جـ	١٤	٢٦	٤٥٠	٦٠٠	٦٣٠٠	١١٧٠٠	٨٤٠٠	١٥٦٠٠
د	٢٤	٤٨	٢٣٠	٤٨٠	٥٥٢٠	١١٠٤٠	١١٥٢٠	٢٣٠٤٠
					٤٥٧٠٠	٨٨٧٤٠	٦٢٣٤٠	١٢٠٥٤٠

$$\text{رقم لاسبير القياسى للأسعار} = \frac{\text{مجـ}\ ع_١ ك_٠}{\text{مجـ}\ ع_٠ ك_٠} \times ١٠٠$$

$$= \frac{٨٨٧٤٠}{٤٥٧٠٠} \times ١٠٠$$

$$= ١,٩٤١٨ \times ١٠٠ = \underline{٪ ١٩٤,١٨}$$

ثانياً : الرقم التجميعى المرجح للأسعار باستخدام كميات سنة المقارنة كأوزان :

الرقم التجميعى للأسعار مرجحاً بكميات سنة المقارنة كأوزان ، يسمى برقم باش القياسى للأسعار The Paasche Price Index .

وإذا رمزنا لهذا الرقم بالرمز $س_١$ ، فإن :

$$س_١ = \frac{\text{مجـ}\ ع_١ ك_١}{\text{مجـ}\ ع_٠ ك_١} \times ١٠٠ \qquad (٧)$$

وهذا يعنى أن رقم باش للأسعار يمثل النسبة المئوية للقيمة الكلية لكميات سنة المقارنة مقومة بأسعار سنة المقارنة إلى القيمة الكلية لنفس هذه الكميات مقومة بأسعار سنة الأساس وبوجه عام . فإن رقم باش للأسعار يحاول الإجابة على السؤال الآتى : « ما هو التغير فى القيمة الكلية لكميات سنة المقارنة عند تقويمها بأسعار سنة الأساس ؟ »

مثــال (٣) :

احسب رقم باش القياسى للأسعار لبيانات الجدول الوارد بالمثال رقم (١) .

الحــل :

باستخدام الأرقام التجميعية المرجحة بالجدول رقم (٢) ، فإن :

$$\text{رقم باش القياسى للأسعار} = \frac{\text{مجـ } ع_1 \, ك_1}{\text{مجـ } ع_. \, ك_1} \times ١٠٠$$

$$= \frac{١٢٠٥٤٠}{٦٢٢٤٠} \times ١٠٠$$

$$= ١,٩٣٣٦ \times ١٠٠ = \underline{١٩٣,٣٦ \%}$$

ثالثاً - الرقم التجميعى المرجح للأسعار باستخدام الوسط الحسابى لكميات سنتى الأساس والمقارنة كأوزان :

بدلاً من استخدام كميات سنة الأساس كأوزان كما هو الحال بالنسبة لرقم لاسبير للأسعار ، أو كميات سنة المقارنة كأوزان كما هو الحال بالنسبة لرقم باش للأسعار . فقد يستخدم الوسط الحسابى لكميات سنتى الأساس والمقارنة كأوزان .

وفى هذه الحالة فإن الأوزان سوف تكون $\frac{ك_. + ك_1}{٢}$ ، وعلى ذلك فإن :

$$\text{الرقم القياسى المطلوب} = \frac{\text{مجـ } ع_1 \left(\frac{ك_. + ك_1}{٢} \right)}{\text{مجـ } ع_. \left(\frac{ك_. + ك_1}{٢} \right)} \times ١٠٠$$

$$= \frac{\text{مجـ } ع_1 (ك_. + ك_1)}{\text{مجـ } ع_. (ك_. + ك_1)} \times ١٠ \qquad (٨)$$

وهذا الرقم يعرف برقم Edgeworth Index

مثال (٤) :

لحساب الرقم القياسى للأسعار المرجح باستخدام الوسط الحسابى لكميات سنتى الأساس والمقارنة كأوزان تنشىء الجدول الآتى :

جدول (٣) حساب رقم إدجو ورث للأسعار

السلعة	الأسعار ٩٠ ع.	الأسعار ٩٥ ع١	الكميات ٩٠ ك.	الكميات ٩٥ ك١	(ك. + ك١)	ع.(ك. + ك١)	ع١(ك. + ك١)
أ	٧٠	١٧٠	٤٠	٣٠	٧٠	٤٩٠٠	١١٩٠٠
ب	٤٢	٨٠	٧٤٠	٩٦٠	١٧٠٠	٧١٤٠٠	١٣٦٠٠٠
جـ	١٤	٢٦	٤٥٠	٦٠٠	١٠٥٠	١٤٧٠٠	٢٧٣٠٠
د	٢٤	٤٨	٢٣٠	٤٨٠	٧١٠	١٧٠٤٠	٣٤٠٨٠
						١٠٨٠٤٠	٢٠٩٢٨٠

بالتعويض فى المعادلة (٨) فإن :

$$\text{رقم إدجو ورث للأسعار} = \frac{٢٠٩٢٨٠}{١٠٨٠٤٠} \times ١٠٠ = ١٩٣,٧١ \%$$

رابعاً : الرقم التجميعي المرجح للأسعار باستخدام الوسط الهندسى لكميات سنتى الأساس والمقارنة كأوزان :

فى هذه الحالة فإن الأوزان سوف تكون $\sqrt{ك_. \, ك_١}$ ، وعلى ذلك فإن :

$$\text{الرقم المطلوب} = \frac{\text{مجـ } ع_١ \sqrt{ك_. \, ك_١}}{\text{مجـ } ع_. \sqrt{ك_. \, ك_١}} \times ١٠٠ \qquad (٩)$$

مثال (٥)

لحساب الرقم القياسي للأسعار مرجحاً باستخدام الوسط الهندسي لكميات سنتي الأساس والمقارنة كأوزان ننشيء الجدول الآتي :

جدول (٤) حساب الرقم التجميعي للأسعار باستخدام الوسط الهندسي للأوزان

السلعة	الأسعار		الكميات					
	٩٠ $ع_٠$	٩٥ $ع_١$	٩٠ $ك_٠$	٩٥ $ك_١$	$ك_٠ \cdot ك_١$	$\sqrt{ك_٠ \cdot ك_١}$	$ع_٠\sqrt{ك_٠ \cdot ك_١}$	$ع_١\sqrt{ك_٠ \cdot ك_١}$
أ	٧٠	١٧٠	٤٠	٣٠	١٢٠٠	٣٤,٦٤	٢٤٢٤,٤٠	٥٨٨٨,٨٠
ب	٤٢	٨٠	٧٤٠	٩٦٠	٧١٠٤٠٠	٨٤٢,٨٥	٣٥٣٩٩,٧٠	٦٧٤٢٨,٠٠
جـ	١٤	٢٦	٤٥٠	٦٠٠	٢٧٠٠٠٠	٥١٩,٦١	٧٢٧٤,٥٤	١٣٥٠٩,٨٦
د	٢٤	٤٨	٢٣٠	٤٨٠	١١٠٤٠٠	٣٣٢,٢٦	٧٩٧٤,٢٤	١٥٩٤٨,٤٨
							٥٣٠٧٣,٢٨	١٠٢٧٧٥,١٤

بالتعويض فى المعادلة (٩) فإن :

الرقم التجميعي المرجح للأسعار باستخدام الوسط الهندسي لكميات سنتي الأساس والمقارنة كأوزان

$$= \frac{١٠٢٧٧٥,١٤}{٥٣٠٧٣,٢٨} \times ١٠٠ = ١٩٣,٦٥\%$$

خامساً : الرقم التجميعي المرجح للأسعار باستخدام الوسط التوافقي لكميات سنتي الأساس والمقارنة كأوزان :

في هذه الحالة فإن الأوزان سوف تكون $= \frac{٢}{\frac{١}{ك_٠} + \frac{١}{ك_١}}$ وعلى ذلك فإن

$$\text{الرقم المطلوب} = \frac{\text{مجـ } ع_{1} \left(\frac{2}{\frac{1}{ك_{1}} + \frac{1}{ك_{0}}} \right)}{\text{مجـ } ع_{0} \left(\frac{2}{\frac{1}{ك_{1}} + \frac{1}{ك_{0}}} \right)} \times 100$$

$$= \frac{\text{مجـ } ع_{1} \left(\frac{1}{\frac{1}{ك_{1}} + \frac{1}{ك_{0}}} \right)}{\text{مجـ } ع_{0} \left(\frac{1}{\frac{1}{ك_{1}} + \frac{1}{ك_{0}}} \right)} \times 100 \qquad (10)$$

وسوف نترك للطالب كتدريب له ، حساب هذا الرقم بتطبيق المعادلة (١٠) على بيانات الجدول الوارد بالمثال رقم (١) .

وسوف يجد الطالب أن حساب هذا الرقم يتطلب عمليات حسابية معقدة وطولة ، لذلك يندر استخدام هذا الرقم في الحياة العملية .

سادساً - الرقم التجميعي المرجح للأسعار باستخدام كميات ثابتة :

إذا رمزنا للأوزان الثابتة بالرمز $ك_{ث}$ ، فإن :

$$\text{الرقم المطلوب} = \frac{\text{مجـ } ع_{1} ك_{ث}}{\text{مجـ } ع_{0} ك_{ث}} \times 100 \qquad (11)$$

مثـال (٦) :

الجدول الآتي يوضح أسعار أربعة سلع أ، ب ، جـ ، د في سنتي ١٩٩٠ ، ١٩٩٥ :

جدول (٥) بيان بأسعار وأوزان بعض السلع

السلعة	الأوزان	الأسعار	
		١٩٩٠	١٩٩٥
أ	١٠	٧٠	١٧٠
ب	٢٥٠	٤٢	٨٠
جـ	١٥٠	١٤	٢٦
د	١٠٠	٢٤	٤٨

والمطلوب عمل الرقم القياسي للأسعار المرجح بالأوزان الواردة بالعمود الثاني عن سنة ١٩٩٥ بالنسبة لسنة ١٩٩٠ كأساس .

الحـــــل :

ننشيء الجدول الآتي :

جدول (٦) حساب الرقم التجميعي للأسعار بأوزان ثابتة .

السلعة	الأوزان	الأسعار ٩٠ ع٠	الأسعار ٩٥ ع١	ع٠ كه	ع١ كه
أ	١٠	٧٠	١٧٠	٧٠٠	١٧٠٠
ب	٢٥٠	٤٢	٨٠	١٠٥٠٠	٢٠٠٠٠
جـ	١٥٠	١٤	٢٦	٢١٠٠	٣٩٠٠
د	١٠٠	٢٤	٤٨	٢٤٠٠	٤٨٠٠
				١٥٧٠٠	٣٠٤٠٠

بالتعويض في المعادلة (١١) فإن :

$$\text{الرقم القياسي المطلوب} = \frac{٣٠٤٠٠}{١٥٧٠٠} \times ١٠٠ = ١٩٣٫٦٣ \%$$

ملاحظات على الأرقام التجميعية للأسعار

الخبرة العملية من استخدام أرقام لاسبير وباش للأسعار تشير إلى أن رقم لاسبير يميل تكبير التغيرات overestimate changes لذلك نجده متحيزاً إلى أعلى upward bias . بينما تشير إلى أن رقم باش يميل إلى تصغير التغيرات (١) underestimate changes لذلك نجده متحيزاً إلى أسفل downward bias وبالرغم من ذلك ، فإن رقم لاسبير ليس بالضرورة أن يكون دائماً أكبر من رقم باش .

وينبغي الإشارة إلى أنه إذا كان من المتعذر التفضيل بين رقمي لاسبير وباش على أسس نظرية ، فإنه يمكن التوفيق بينهما على أسس تطبيقية وذلك إما بأخذ الوسط الحسابي لهما أو الوسط الهندسي لهما .

ففي عام ١٨٧١ ، اقترح درويش Drobisch أخذ الوسط الحسابي لرقمي لاسبير وباش . وهذا يعني أن :

$$\text{رقم درويش القياسي للأسعار} = \frac{\dfrac{\sum \text{ع}_1 \text{ك}_0}{\sum \text{ع}_0 \text{ك}_0} + \dfrac{\sum \text{ع}_1 \text{ك}_1}{\sum \text{ع}_0 \text{ك}_1}}{٢} \times ١٠٠ \quad (١٢)$$

أما في عام ١٩٢٠ ، فقد اقترح ارفنج فيشر Irving Fisher رقماً قياسياً يقع بين رقمي لاسبير وباش القياسيين ، وذلك بأخذ الوسط الهندسي لهما . وهذا الرقم القياسي يعرف « بالرقم القياسي الأمثل » Ideal Index . والسبب في تسميته بهذا الإسم هو أنه يحقق الخصائص الرياضية لتكوين الرقم القياسي الجيد ، وهي قابليته للانعكاس في الزمن وللانعكاس في المعامل على الوجه الذي سيتم شرحه تفصيلاً فيما بعد .

---

(١) بالنسبة للسلع التي ارتفعت اسعارها ، فإن المستهلكين يقللون الكميات المستهلكة منها . ومن ثم تكون الكميات المستهلكة منها في فترة المقارنة اقل من الكميات المستهلكة منها في فترة الأساس ونظراً لان رقم لاسبير يستخدم كميات سنة الأساس كأوزان في بسط ومقام المعادلة الخاصة بحسابة ، فإنه بذلك يعطي وزناً أكبر للسلع التي ارتفعت اسعارها مما يؤدي الى تكبير بسط المعادلة وبالتالي يؤدي الى تكبير قيمة هذا الرقم القياسي . أما رقم باش فيستخدم كميات سنة المقارنة كأوزان في بسط ومقام المعادلة الخاصة بحسابه . مما يؤدي الى تصغير التغيرات وبالتالي الي تصغير قيمة هذا الرقم

وعلى ذلك فإن :

$$\text{رقم فيشر الأمثل} = \sqrt{\frac{\text{مجـ}\ ع_{١}\ ك_{.}}{\text{مجـ}\ ع_{.}\ ك_{.}} \times \frac{\text{مجـ}\ ع\ ك_{١}}{\text{مجـ}\ ع_{.}\ ك_{١}}} \times ١٠٠ \qquad (١٢)$$

وحيث أننا حسبنا رقم لاسبير القياسى للأسعار لبيانات الجدول الوارد بالمثال (١) وذلك فى المثال (٢) ووجدنا أن قيمته = ١٩٤،١٨ . وأيضا حسبنا رقم باش القياسى للأسعار لنفس بيانات هذا الجدول فى المثال (٣) ووجدنا أن قيمته = ١٩٣،٣٦ وعلي ذلك فإن :

$$\text{رقم فيشر الأمثل} = \sqrt{١٩٤،١٨ \times ١٩٣،٣٦}$$

$$= \sqrt{٣٧٥٤٦،٦٤} = ١٩٣،٧٧ \%$$

ثانياً : بالنسبة للكميات

الأرقام القياسية التجميعية المرجحة للكميات هى الأرقام المقابلة للأرقام القياسية التجميعية المرجحة للأسعار . فبينما كنا نستخدم الكميات كأوزان عند حساب الأرقام التجميعية للأسعار ، فإنه فى حالتنا هذه نستخدم الأسعار كأوزان .

وهذا وللحصول على المعادلات الخاصة بحساب الأرقام التجميعية المرجحة للكميات، فإن كل ما يستلزمه الأمر هو استبدال الرمز ع بالرمز ك ، والرمز ك بالرمز ع مع بقاء الأدلة السفلية على حالها فى المعادلات المناظرة لها فى الأرقام القياسية التجميعية المرجحة للأسعار .

وأهم الأرقام التجميعية المرجحة للكميات هما الرقمين الآتيين :

الرقم الأول : الرقم التجميعى للكميات مرجحاً بأسعار سنة الأساس كأوزان .

وهو ما يعرف برقم لاسبير للكميات ، وسوف نرمز له بالرمز ك_{0} وعلى ذلك فإن :

$$\text{ك}_{0} = \frac{\text{مجـ } \text{ك}_{1}\,\text{ع}_{0}}{\text{مجـ } \text{ك}_{0}\,\text{ع}_{0}} \times 100 \qquad (14)$$

الرقم الثانى : الرقم التجميعى للكميات مرجحاً بأسعار سنة المقارنة كأوزان .

وهو ما يعرف برقم باش للكميات ، وسوف نرمز له بالرمز ك_{1} وعلى ذلك فإن :

$$\text{ك}_{1} = \frac{\text{مجـ } \text{ك}_{1}\,\text{ع}_{1}}{\text{مجـ } \text{ك}_{0}\,\text{ع}_{1}} \times 100 \qquad (15)$$

مثال (٧) :

لحساب رقم لاسبير القياسى للكميات ، ورقم باش القياسى للكميات لبيانات الجدول الوارد بالمثال رقم (١) فإننا نستخدم الأرقام التجميعية الواردة في الجدول رقم (٣) الخاص بالمثال رقم (٢)

وعلى ذلك فإن رقم لاسبير للكميات هو :

$$ك = \frac{مجـ ك_1 ع_.}{مجـ ك_. ع_.} \times ١٠٠$$

$$= \frac{٦٢٢٤٠}{٤٥٧٠٠} \times ١٠٠$$

$$= ١,٣٦٤١ \times ١٠٠ = ١٣٦,٤١\%$$

ورقم لاسبير للكميات يعنى أنه : باستخدام أسعار سنة ١٩٩٠ كأساس ، فإن حجم الإنتاج the volumes of output يكون قد زاد بنسبة ٣٦,٤١ ٪ بين سنتى ١٩٩٠ ، ١٩٩٥ .

أما رقم باش للكميات فيكون :

$$ك_1 = \frac{مجـ ك_1 ع_1}{مجـ ك_. ع_1} \times ١٠٠$$

$$= \frac{١٢٠٥٤٠}{٨٨٧٤٠} \times ١٠٠$$

$$= ١,٣٥٨٤ \times ١٠٠ = ١٣٥,٨٤ \%$$

ورقم باش للكميات يعنى أنه : باستخدام أسعار سنة ١٩٩٥ كأساس .
فإن حجم الانتاج يكون قد زاد بنسبة ٣٥,٨٤٪ بين سنتى ١٩٩٠، ١٩٩٥ .

وبصفة عامة ، فإن الرقم التجميعى المرجح للكميات يجيب على السؤال الآتى :

« إذا قمنا بشراء (أو ببيع) كميات مختلفة من نفس السلع فى كل من سنتى المقارنة والأساس بنفس السعر at the same price ، فكم ننفق (أو نتسلم) فى فترة المقارنة بالنسبة لفترة الأساس ؟ »

الفصل الثانى
بعض موضوعات خاصة بالارقام القياسية

تغيير الاساس
Changing The Base Period

عادة ما يتطلب الأمر تغيير أساس سلسلة من الأرقام القياسية من فترة إلى أخرى وذلك دون الرجوع إلى البيانات الأصلية ، وإعادة حساب السلسلة بأكملها على نظام الأساس الجديد . وغالباً ما يرجع ذلك إلى إحدى السببين الآتين .

أولا: إذا كانت فترة الأساس التى ينسب إليها الرقم القياسى بعيدة جداً عن فترة المقارنة الأمر الذى يؤثر تأثيراً كبيراً على الرقم القياسى ويفقده صفة التمثيل . لذلك فإن الأمر يستلزم تغيير فترة الأساس إلى فترة قريبة من فترة المقارنة ، حتى تكون متمشية مع الظروف الاقتصادية المحيطة بالظاهرة موضوع الدراسة .

ثانياً: إذا كان لدينا سلسلتين من الأرقام القياسية ، وأساس كل منهما مختلف عن الأخرى . وإذا أريد إجراء مقارنة بين هاتين السلسلتين . فمن البديهى أنه لا يمكن إجراء المقارنة المطلوبة إلا بتوحيد فترة الأساس بالنسبة لكل منهما . ويتم ذلك بتغيير فترة الأساس بالنسبة لإحداها إلى فترة الأساس بالنسبة إلى الأخرى حتى يتسنى إجراء عملية المقارنة .

هذا ، ويعد تغييرأساس الرقم القياسى أمراً فى غاية البساطة ، فعلى سبيل المثال ، إذا كان لدينا سلسلة من الأرقام القياسية مركبة على اعتبار أن سنة ١٩٨٨ كأساس . وإذا أردنا تغيير سنة الأساس إلى سنة حديثة ولتكن سنة ١٩٩٢ ، فإن الأمر لا يستلزم سوى قسمة كل رقم من أرقام السلسلة القديمة على الرقم القياسى لسنة ١٩٩٢ ثم ضرب خارج القسمة فى ١٠٠.

مثال :

الجدول الآتى يوضح تغيير أساس سلسلة الأرقام القياسية من سنة ١٩٨٨ كأساس إلى سنة ١٩٩٢ كأساس :

جدول (١) تغيير الأساس

السنـــة	الرقم القياسي الأصلي ١٩٨٨ = ١٠٠	الرقم القياسي المعدل ١٩٩٢ = ١٠٠
١٩٨٧	٩٨,٤	٨٦,٧
١٩٨٨	١٠٠,٠	٨٨,١
١٩٨٩	١٠٤,٧	٩٢,٢
١٩٩٠	١٠٦,٢	٩٣,٥
١٩٩١	١٠٩,٤	٩٦,٤
١٩٩٢	١١٣,٥	١٠٠,٠
١٩٩٣	١١٨,٨	١٠٤,٧
١٩٩٤	١٢٣,٢	١٠٨,٥
١٩٩٥	١٢٩,١	١١٣,٧

وعلى ذلك فإن :

الرقم القياسي لسنة ١٩٨٧ على اعتبار أن سنة ١٩٩٢ كأساس

$$= \frac{٩٨,٤}{١١٣,٥} \times ١٠٠ = ٨٦,٧$$

، الرقم القياسي لسنة ١٩٨٨ على اعتبار أن سنة ١٩٩٢ كأساس

$$= \frac{١٠٠,٠}{١١٣,٥} \times ١٠٠ = ٨٨,١$$

وهكذا بالنسبة لسائر السنوات .

وتجدر الإشارة إلى أن العلاقة بين سلسلة الأرقام القياسية الجديدة بعد تغيير فترة الأساس إلى سنة ١٩٩٢ ، هي نفس العلاقة بين سلسلة الأرقام القياسية الأصلية . فعلى سبيل المثال ، فإننا نجد أن الرقم القياسي لعام ١٩٨٨ يزيد على الرقم القياسي لعام ١٩٨٧ بنفس النسبة في السلسلتين . وهذا يعني أن:

$$\frac{١٠٠}{٩٨,٤} = \frac{٨٨,١}{٨٦,٧} = ١,٠١٦$$

وبالمثل ، فإن الرقم القياسي لعام١٩٨٩ يزيد على الرقم القياسي لعام ١٩٨٨ بنفس النسبة في السلسلتين . وهذا يعني أن .

$$\frac{١٠٤,٧}{١٠٠} = \frac{٩٢,٢}{٨٨,١} = ١,٠٤٧$$

وهكذا بالنسبة لسائر أرقام السلسلتين .

* * *

الفصل الثالث

اختبارات الأرقام القياسية
Tests of Index Numbers

تمهيد :

رأينا أن هناك صيغاً عديدة تستخدم لحساب الأرقام القياسية من نفس بيانات الأسعار ومن نفس بيانات الكميات . وكل منها عند استخدامها يعطينا نتائج مختلفة . والسؤال الذى يواجهنا الآن هو : « أى من هذه الصيغ يعتبر أكثر دقة ؟ » .

طبقاً للإحصائيين الرياضيين ، فإن مدى الدقة المتوقعة للصيغة المستخدمة يتوقف على مدى قابليتها لاجتياز اختبارات رياضية معينة . ومن هذه الاختبارات ، وربما يكون أكثرها أهمية فى مجال الأرقام القياسية ، هى الاختبارات :

الأول : اختبار الانعكاس فى الزمن .

الثانى : اختبار الانعكاس فى العامل .

وسوف نتناول شرح كل منها بشىء من التفصيل فى هذا الفصل .

أولا : اختبار الانعكاس فى الزمن
The Time Reversal Test

يقضى هذا الاختبار بأنه إذا كان الرقم القياسى للفترة « ٠ » (أى فترة الأساس) بالنسبة للفترة « ١ » (أى فترة المقارنة) كأساس ، عبارة عن مقلوب reciprocal الرقم القياسى للفترة « ١ » بالنسبة للفترة « ٠ » كأساس ، فإن حاصل ضرب الرقمين يكون مساوياً للواحد الصحيح[1] . وفى هذه الحالة يقال أن الرقم القياسى قد اجتاز شرط الانعكاس الزمن .

---

(١) وهذا أمر بديهى ، وذلك لأن حاصل ضرب أى نسبة فى مقلوبها لابد وأن تكون مساوية للواحد الصحيح

وعلى سبيل المثال ، فسوف نأخذ المثال المبسط الآتى :

إذا فرضنا أن سعر السلعة أ فى سنة ١٩٩٥ هو ٣٠ قرشاً ، وكان سعرها فى سنة ١٩٩٠ هو عشرة قروش . فإن الرقم القياسى لسعر هذه السلعة فى سنة ١٩٩٥ بالنسبة إلى سنة ١٩٩٠ كأساس هو ٣٠٠ ٪ ، بينما يكون الرقم القياسى لسعر هذه السلعة فى سنة ١٩٩٠ بالنسبة إلى سنة ١٩٩٥ كأساس هو ٣٣ $\frac{1}{3}$ ٪ . وهذا أمر منطقى ، لأنه طالما أن مستوى السعر بالنسبة إلى هذه السلعة فى سنة ١٩٩٥ قد بلغ ثلاثة أمثال ما كان عليه فى سنة ١٩٩٠، فإن مستوى السعر فى سنة ١٩٩٠ بالنسبة إلى سنة ١٩٩٥ كأساس لابد وأن يكون مساوياً للثلث .

وعلى ضوء ما تقدم ، فإنه لحساب البديل الزمنى لأى صيغة من صيغ معادلات الأرقام القياسية السابق دراستها ، فإن الأمر يستلزم تغيير الأدلة السفلية الملحقة برموز الأسعار (ع) أو الكميات (ك) من الدليل « ٠ » إلى الدليل « ١ » ومن الدليل « ١ » إلى الدليل « ٠ » . أو بعبارة أخرى نستبدل أسعار سنة الأساس وكميات سنة الأساس بأسعار سنة المقارنة وكميات سنة المقارنة ، وبالعكس من ذلك نستبدل أسعار سنة المقارنة وكميات سنة المقارنة بأسعار سنة الأساس وكميات سنة الأساس . وبإجراء ذلك، نحصل على البديل الزمنى لصيغة الرقم القياسى المراد اختباره . ويكون هذا الرقم قابلا للانعكاس فى الزمن إذا حقق الشرط الآتى :

الرقم القياسى × بديله الزمنى = واحداً صحيحاً .

مثال (١) :

اختبر ما إذا كانت الأرقام القياسية الآتية تحقق شرط الانعكاس فى الزمن من عدمه مع إيضاح أسباب ذلك :

أولا : الرقم التجميعى البسيط للأسعار .

ثانياً : رقم لاسبير القياسى للأسعار .

ثالثاً : رقم باش القياسى للأسعار .

الحل .

فى عملية الاختبار المطلوبة ، سوف نقوم بإهمال العامل (١) من المعادلات المختصة

أولاً :الرقم التجميعى البسيط للأسعار :

من المعادلة (٤) نجد أن : صيغته = $\frac{\text{مجـ ع}_1}{\text{مجـ ع}_0}$

∴ بديله الزمنى = $\frac{\text{مجـ ع}_0}{\text{مجـ ع}_1}$

حاصل الضرب = $\frac{\text{مجـ ع}_1}{\text{مجـ ع}_0} \times \frac{\text{مجـ ع}_0}{\text{مجـ ع}_1} = 1$

∴ الرقم التجميعى البسيط للأسعار يقبل الانعكاس فى الزمن .

ثانياً : رقم لاسبير القياسى للأسعار :

من المعادلة (٦) نجد أن : صيغته = $\frac{\text{مجـ ع}_1 \text{ك}_0}{\text{مجـ ع}_0 \text{ك}_0}$

بديله الزمنى = $\frac{\text{مجـ ع}_0 \text{ك}_1}{\text{مجـ ع}_1 \text{ك}_1}$

حاصل الضرب = $\frac{\text{مجـ ع}_1 \text{ك}_0}{\text{مجـ ع}_0 \text{ك}_0} \times \frac{\text{مجـ ع}_0 \text{ك}_1}{\text{مجـ ع}_1 \text{ك}_1} \neq 1$.

∴ رقم لا سبير القياسى للأسعار لايجتاز اختبار الانعكاس فى الزمن .

ثالثاً : رقم باش القياسى للأسعار :

من المعادلة (٧) نجد أن : صيغته = $\frac{\text{مجـ ع}_1 \text{ك}_1}{\text{مجـ ع}_0 \text{ك}_1}$

بديله الزمنى = $\frac{\text{مجـ ع}_0 \text{ك}_0}{\text{مجـ ع}_1 \text{ك}_0}$

حاصل الضرب = $\frac{\text{مجـ ع}_1 \text{ك}_1}{\text{مجـ ع}_0 \text{ك}_1} \times \frac{\text{مجـ ع}_0 \text{ك}_0}{\text{مجـ ع}_1 \text{ك}_0} \neq 1$

∴ رقم باش القياسى للأسعار لايجتاز اختبار الانعكاس فى الزمن .

ثانياً: اختبار الانعكاس فى المعامل
The Factor Reversal Test

يقضى هذا الاختبار بأن : حاصل ضرب الرقم القياسى للأسعار فى الرقم القياسى للكميات ينبغى أن يكون مساوياً للرقم القياسى المناظر للرقم القياسى للقيمة . وقد سبق أن تكلمنا عن هذه الخاصية بشيء من التفصيل عند التكلم عن « الأرقام القياسية للقيمة والاتساق بين الأرقام القياسية للأسعار والكميات »

وما نود الإشارة إليه هنا هو أن :

أولا : إذا كان حاصل ضرب الرقم القياسى × بديله المعاملى مساوياً للرقم القياسى للقيمة ، فإنه يقال أن الرقم القياسى قد اجتاز شرط الانعكاس فى المعامل .

ويتم الحصول على البديل المعاملى لصيغة الرقم القياسى المراد اختباره : وذلك ، باستبدال رموز الأسعار (ع) برموز الكميات (ك) ، وبالعكس من ذلك استبدال رموز الكميات (ك) برموز الأسعار (ع) مع بقاء الأدلة السفلية الملحقة بالرموز فى صيغة معادلة الرقم القياسى الأصلية على حالها .

ثانياً : أن جميع صيغ الأرقام القياسية السابق دراستها - وذلك باستثناء رقم فيشر الأمثل - غير قابلة لاجتياز شرط الانعكاس فى المعامل .

مثال (٢) :

اختبرما إذا كانت الأرقام القياسية الآتية تحقق شرط الانعكاس فى المعامل من عدمه ، مع إيضاح أسباب ذلك :

أولا : الرقم التجميعى البسيط للأسعار .

ثانياً : رقم لاسبير القياسى للأسعار .

ثالثاً : رقم باش القياسى للأسعار .

الحل :

فى عملية ألاختبار المطلوبة ، سوف نقوم بإهمال العامل (١٠٠) من المعادلات المختصة .

أولا : الرقم التجميعى البسيط للأسعار :

من المعادلة (٤) نجد أن : صيغته = $\frac{\text{مجـ ع}_1}{\text{مجـ ع}_0}$

بديله العاملى = $\frac{\text{مجـ ك}_1}{\text{مجـ ك}_0}$

$$\text{حاصل الضرب} = \frac{\text{مجـ ع}_1}{\text{مجـ ع}_0} \times \frac{\text{مجـ ك}_1}{\text{مجـ ك}_0} \neq \frac{\text{مجـ ع}_1\text{ك}_1}{\text{مجـ ع}_0\text{ك}_0}$$

وهذا يعنى أن الرقم التجميعى البسيط للأسعار لايقبل الانعكاس فى العامل .

ثانياً : رقم لاسبير القياسى للأسعار :

من المعادلة (٦) نجد أن : صيغته = $\frac{\text{مجـ ع}_1\text{ك}_0}{\text{مجـ ع}_0\text{ك}_0}$

بديله العاملى = $\frac{\text{مجـ ك}_1\text{ع}_0}{\text{مجـ ك}_0\text{ع}_0}$

$$\text{حاصل الضرب} = \frac{\text{مجـ ع}_1\text{ك}_0}{\text{مجـ ع}_0\text{ك}_0} \times \frac{\text{مجـ ك}_1\text{ع}_0}{\text{مجـ ك}_0\text{ع}_0} \neq \frac{\text{مجـ ع}_1\text{ك}_1}{\text{مجـ ع}_0\text{ك}_0}$$

وهذا يعنى أن رقم لاسبير القياسى للأسعار لا يقبل الانعكاس فى العامل .

ثالثاً : رقم باش القياسى للأسعار :

من المعادلة (٧) نجد أن صيغته = $\frac{\text{مجـ ع}_1\text{ك}_1}{\text{مجـ ع}_0\text{ك}_1}$

بديله العاملى = $\frac{\text{مجـ ك}_1\text{ع}_1}{\text{مجـ ك ع}_1}$

$$\text{حاصل الضرب} = \frac{\text{مجـ ع}_1\text{ك}_1}{\text{مجـ ع}_0\text{ك}_1} \times \frac{\text{مجـ ك}_1\text{ع}_1}{\text{مجـ ك ع}_1} \neq \frac{\text{مجـ ع}_1\text{ك}_1}{\text{مجـ ع ك}}$$

وهذا يعنى أن رقم باش القياسى للأسعار لا يقبل الانعكاس فى العامل .

ملاحظات على اختبارات الأرقام القياسية ورقم فيشر الأمثل:

رأينا أن غالبية الأرقام القياسية الشائعة الاستخدام وخاصة الأرقام القياسية المرجحة غير قابلة لاجتياز معظم الاختبارات السابقة. وما نود الإشارة إليه هو أن ذلك لا يعنى بالضرورة التقليل أو الإضرار بأهميتها ، كما أن ذلك لا يبرهن على أنه ليس لهذه الأرقام خصائص منطيقية مفيدة . فقد سبق أن رأينا عند دراستنا لأرقام لاسبير وباش القياسيين أنهما – على الرغم من عدم دقتهما من وجهة النظر الرياضية المنطقية – يعطيان إجابات مفيدة لأنواع خاصة من الأسئلة الهامة والمراد الحصول على إجابات عنها .

وهناك عدة محاولات قد بذلت بمعرفة ارفنج فيشر لتكوين صيغاً من الأرقام القياسية لكى تجتاز معظم هذه الاختبارات الرياضية . ومن هذه الصيغ الهامة والسابق الإشارة إليها هو « رقم فيشر الأمثل » الذى هوعبارة عن الوسط الهندسى لحاصل ضرب رقمى لاسبير وباش القياسيين للأسعار ، ومعادلته هى رقم (١٣) .

وقد سبق أن ذكرنا أن السبب فى تسمية رقم فيشر بالرقم القياسى الأمثل هى قابليته للانعكاس فى الزمن و للانعكاس فى المعامل على الوجه الذى سنوضحه فيما يلى :

أولا - اختبار الانعكاس فى الزمن :

$$\text{صيغة رقم فيشر} = \sqrt{\frac{\text{مجـ ع}_1\text{ك}_0}{\text{مجـ ع}_0\text{ك}_0} \times \frac{\text{مجـ ع}_1\text{ك}_1}{\text{مجـ ع}_0\text{ك}_1}}$$

معادلة رقم (١٣) بعد إهمال العامل ١٠٠ .

وللحصول على بديله الزمنى نقوم - كما سبق أن ذكرنا - باستبدال الدليل « ٠ » بالدليل « ١ » وكذا الدليل « ١ » بالدليل « ٠ » على النحو الآتى :

$$\text{البديل الزمنى} = \sqrt{\frac{\text{مجـ ع}_0\text{ك}_1}{\text{مجـ ع}_1\text{ك}_1} \times \frac{\text{مجـ ع}_0\text{ك}_0}{\text{مجـ ع}_1\text{ك}_0}}$$

$$\text{حاصل الضرب} = \sqrt{\frac{\text{مجـ ع}_1\text{ك}_0}{\text{مجـ ع}_0\text{ك}_0} \times \frac{\text{مجـ ع}_1\text{ك}_1}{\text{مجـ ع}_0\text{ك}_1}} \times \sqrt{\frac{\text{مجـ ع}_0\text{ك}_1}{\text{مجـ ع}_1\text{ك}_1} \times \frac{\text{مجـ ع}_0\text{ك}_0}{\text{مجـ ع}_1\text{ك}_0}}$$

$$= \sqrt{\frac{\text{مجـ ع}_1\text{ك}_0}{\text{مجـ ع}_0\text{ك}_0} \times \frac{\text{مجـ ع}_1\text{ك}_1}{\text{مجـ ع}_0\text{ك}_1} \times \frac{\text{مجـ ع}_0\text{ك}_1}{\text{مجـ ع}_1\text{ك}_1} \times \frac{\text{مجـ ع}_0\text{ك}_0}{\text{مجـ ع}_1\text{ك}_0}}$$

$$= 1$$

وهذا يعنى أن رقم فيشر القياسى يقبل شرط الانعكاس فى الزمن، وهذا يمثل السبب الأول فى تسمية بالرقم القياسى الأمثل .

ثانياً - اختبار الانعكاس فى المعامل :

$$\text{صيغة رقم فيشر القياسى} = \sqrt{\frac{\text{مجـ ع}_1\text{ك}_0}{\text{مجـ ع}_0\text{ك}_0} \times \frac{\text{مجـ ع}_1\text{ك}_1}{\text{مجـ ع}_0\text{ك}_1}}$$

(بإهمال العامل ١)

وللحصول على بديله المعاملى نقوم - كما سبق أن ذكرنا - باستبدال رموز الأسعار (ع) برموز الكميات (ك) ، وكذا استبدال رموز الكميات (ك) برموز الأسعار (ع) ، مع بقاء الأدلة السفلية للرموز فى الصيغة السابقة على ما هى عليه وعلى ذلك فإن

البديل العاملي = $\sqrt{\frac{\text{مجـ ك}_1\text{ ع}_0}{\text{مجـ ك}_0\text{ ع}_0} \times \frac{\text{مجـ ك}_1\text{ ع}_1}{\text{مجـ ك}_0\text{ ع}_1}}$

حاصل الضرب = $\sqrt{\frac{\text{مجـ ع}_1\text{ ك}_0}{\text{مجـ ع}_0\text{ ك}_0} \times \frac{\text{مجـ ع}_1\text{ ك}_1}{\text{مجـ ع}_0\text{ ك}_1}} \times \sqrt{\frac{\text{مجـ ك}_1\text{ ع}_0}{\text{مجـ ك}_0\text{ ع}_0} \times \frac{\text{مجـ ك}_1\text{ ع}_1}{\text{مجـ ك}_0\text{ ع}_1}}$

$= \sqrt{\frac{\text{مجـ ع}_1\text{ ك}_0}{\text{مجـ ع}_0\text{ ك}_0} \times \frac{\text{مجـ ع}_1\text{ ك}_1}{\text{مجـ ع}_0\text{ ك}_1} \times \frac{\text{مجـ ك}_1\text{ ع}_0}{\text{مجـ ك}_0\text{ ع}_0} \times \frac{\text{مجـ ك}_1\text{ ع}_1}{\text{مجـ ك}_0\text{ ع}_1}}$

$= \sqrt{\frac{\text{مجـ ع}_1\text{ ك}_1}{\text{مجـ ع}_0\text{ ك}_0} \times \frac{\text{مجـ ع}_1\text{ ك}_1}{\text{مجـ ع}_0\text{ ك}_0}} = \frac{\text{مجـ ع}_1\text{ ك}_1}{\text{مجـ ع}_0\text{ ك}_0}$

= ق (الرقم القياسي للقيمة)

وهذا يعني أن رقم فيشر القياسي يقبل شرط الانعكاس في العامل ، وهذا يمثل السبب الثاني في تسميته بالرقم القياسي الأمثل .

هذا ، وهناك عدة انتقادات توجه إلى رقم فيشر القياسي الأمثل ، نلخص أهمها فيما يلي :

أولا : صعوبة الحساب .

ثانياً : استخدام كميات سنة المقارنة في الترجيح يتطلب تكاليف باهظة فضلا عن الوقت والجهد المطلوبين .

ثالثاً : من الصعوبة بمكان إعطاء تفسيرات دقيقة لما يحاول هذا الرقم حقيقة قياسه .

* * *

مثال (٣) :

وضح أن رقم فيشر القياسي الأمثل لسنة ١٩٩٥ بالنسبة إلى سنة ١٩٩٠ كأساس ، يحقق اختباري الانعكاس في الزمن والانعكاس في العامل ، وذلك بالنسبة لبيانات الجدول الوارد بالمثال (٢) الحـل :

سبق أن حسبنا الآتي :

رقم لاسبير القياسي للأسعار ووجدناه = ١٩٤.١٨ ٪

رقم باش القياسي للأسعار ووجدناه = ١٩٣.٣٦ ٪

رقم فيشر " " ووجدناه = ١٩٣ ٧٧ ٪

وسنجرى الاختبارين المطلوبين على الترتيب الآتى :

أولا : اختبار الانعكاس فى الزمن :

بالاستعانة ببيانات جدول (٣) الخاص بالأرقام التجميعية المرجحة للمثال المذكور ، فانه بالتعويض فى معادلة البديل الزمنى لرقم فيشر نجد أن :

البديل الزمنى = √(٦٢٢٤٠ / ١٢٠٥٤٠ × ٤٥٧٠٠ / ٨٨٧٤٠)

= √(٠,٥١٧١٧٣ × ٠,٥١٤٩٨٨)

= √ ٠,٢٦٦٢٢٨ = **٠,٥١٦**

ولتحقيق اختبار الانعكاس فى الزمن ينبغى أن يكون :

رقم فيشر القياسى × بديله الزمنى = واحداً صحيحاً .

حاصل الضرب = ١,٩٣٧٧ × ٠,٥١٦ = ٠,٩٩٩٩ = واحداً صحيحاً .

وهذا يعنى أن رقم فيشر الأمثل يقبل الانعكاس فى الزمن .

ثانياً : اختبار الانعكاس فى المعامل :

بالتعويض فى البديل المعاملى لرقم فيشر من بيانات جدول (٢) :

البديل المعاملى لرقم فيشر = √(٦٢٢٤٠ / ٤٥٧٠٠ × ١٢٠٥٤٠ / ٨٨٧٤٠)

= √(١,٣٦٤١١٤ × ١,٣٥٨٣٥٠)

= √ ١,٨٥٢٩٤٥ = **١,٣٦٢٧**

. نحسب الرقم القياسي للقيمة حيث = $\frac{\text{مجـ ع}_1 \text{ ك}_1}{\text{مجـ ع}_0 \text{ ك}_0}$

$$= \frac{١٢٠٥٤٠}{٤٥٧٠٠} = ٢,٦٣٧٦$$

ولتحقيق اختبار الانعكاس في المعامل ينبغي أن يكون :

رقم فيشر القياسي × بديله المعاملي = الرقم القياسي للقيمة

حاصل الضرب = ١,٩٣٧٧ × ١,٣٦١٢

= <u>٢,٦٣٧٦</u> = الرقم القياسي للقيمة

وهذا يعني أن رقم فيشر الأمثل يقبل الانعكاس في المعامل .

تطبيقات الأرقام القياسية

١ - اشرح بالتفصيل المقصود بالرقم القياسي وكيفية استخدامه للدلالة على التغيرات الاقتصادية والتجارية .

٢ - تكلم عن مزايا وعيوب كل من رقمى لاسبير وباش القياسيين للأسعار ، مع ذكر الأسباب الرئيسية التى تؤدى إلى زيادة الاختلافات بين هذين الرقمين .

٣ - عرف رقم فيشر القياسى للأسعار ، مع إيضاح أسباب تسمية بالرقم القياسى الأمثل .

٤ - تستخدم فى الأرقام القياسية عدة اختبارات لاختبار جودة الرقم القياسى .
اشرح ماهية هذه الاختبارات مع إيضاح مدى انطباقها على أرقام لاسبير وباش وادجورث القياسية للأسعار .

٥ - البيانات الآتية توضح الأسعار بالجنيه والكميات بآلاف الوحدات لثلاثة أنواع من السلع أ ، ب ، جـ عن عامي ١٩٩٢ ، ١٩٩٥ :

السلعة	١٩٩٢		١٩٩٥	
	السعر	الكمية	السعر	الكمية
أ	١٠	١٢	١٣	١٥
ب	٢٠	٤	٢٦	٦
جـ	٢٥	٨	٣٠	١٠

والمطلوب حساب الأرقام القياسية الآتية لسنة ١٩٩٥ بالنسبة إلى سنة ١٩٩٢ كأساس باستخدام :

أولا : الرقم التجميعى البسيط للأسعار .

ثانياً : الرقم التجميعى المرجح للأسعار باستخدام كميات سنة الأساس كأوزان .

ثالثاً : الرقم التجميعى المرجح للأسعار باستخدام كميات سنة المقارنة كأوزان .

رابعاً :الرقم التجميعى المرجح للأسعار باستخدام الوسط الحسابى لكميات سنتى الأساس والمقارنة كأوزان .

خامساً : رقم فيشر القياسى الأمثل .

٦ - الجدول الآتى يوضح متوسط الأجور الأسبوعية بالجنيهات لأربعة صناعات أ ، ب ، جـ ، د وعدد العمال فى كل منها بالنسبة لشهر يناير من سنتى ١٩٩٤ ، ١٩٩٥

الصناعة	عدد العمال	متوسط الأجور الأسبوعية بالجنيهات	
		١٩٩٤	١٩٩٥
أ	٧٥٠٠	٤٢	٥٦
ب	٤٦٠٠	٣٦	٤٥
هـ	٣٠٠٠	٣٢	٣٨
د	١٢٠٠	٢٥	٣٠

والمطلوب عمل رقم قياسي للأجور لسنة ١٩٩٥ بالنسبة الى سنة ١٩٩٤ كأساس .

٧- الجدول الآتى يوضح متوسط الأجور الأسبوعية بالجنيهات لأربعة محافظات أ ، ب ، جـ ، د وعدد العمال فى كل منها بالنسبة لشهر يناير من سنتى ١٩٩٠ ، ١٩٩٥ :

المحافظات	عدد العمال		متوسط الأجور الأسبوعية بالجنيهات	
	١٩٩٠	١٩٩٥	١٩٩٠	١٩٩٥
أ	٧٢٠٠	١١٦٥٠٠	٧٠	١٠٠
ب	٤٢٥٠٠	٦٤٥٠٠	٦٠	٨٠
هـ	٣٦٤٠٠	٥٢٢٠٠	٥٠	٧٠
د	٢٤٥٠٠	٣٤٢٥٠	٣٠	٤٠

والمطلوب عمل رقم قياسى للأجور لسنة ١٩٩٥ بالنسبة إلى سنة ١٩٩٠ كأساس

٨ - الجدول الآتى يوضح سلسلة الأرقام القياسية من سنة ١٩٨٨ إلى ١٩٩٥ .
وقد اتخذت سنة ١٩٨٩ كأساس

السنة	١٩٨٨	١٩٨٩	١٩٩٠	١٩٩١	١٩٩٢	١٩٩٣	١٩٩٤	١٩٩٥
الرقم القياسى ١٩٨٩ = ١٠٠	٩٦,٨	١٠٠,٠	١٠٥,٨	١١٢,٢	١٢٠,٨	١٣٢,٥	١٤٨,٦	١٦٤,٢

والمطلوب تغيير أساس هذه السلسلة إلى سنة ١٩٩١ .

٩- الجدول الآتى يوضح أسعار وكميات أربعة سلع أ ، ب ، جـ ، د :

السلعة	الكميات	الأسعار	
		١٩٩٠	١٩٩٥
أ	٥٠	٥	٨
ب	١٠٠	٧	١٢
هـ	٢٠٠	١٠	١٦
د	٥٠٠	١٥	٢٤

والمطلوب عمل الرقم القياسى المناسب للأسعار لسنة ١٩٩٥ بالنسبة إلى سنة ١٩٩٠ كأساس .

١٠- البيانات الآتية توضح الأسعار بالجنيه والكميات بالآلاف لثلاثة سلع أ ، ب ، جـ فى السنتين ١٩٩٠ ، ١٩٩٥ :

السلعة	الأسعار		الكميات	
	١٩٩٠	١٩٩٥	١٩٩٠	١٩٩٥
أ	١٥	١٨	٢٠	٤٠
ب	٨	١٢	٦٠	١٠٠
هـ	٧	٨	٢٠٠	٣٠٠

والمطلوب حساب الأرقام القياسية الآتية لسنة ١٩٩٥ بالنسبة إلى سنة ١٩٩٠ كأساس باستخدام :

أولاً : رقم لاسبير القياسى للأسعار ،

ثانيـاً : رقم باش القياسى للأسعار .

ثالثـاً : رقم ادجوورث القياسى للأسعار .

رابعـاً : رقم فيشر الأمثل .

خامساً : إيضاح أن رقم فيشر الأمثل يحقق اختبار الانعكاس فى الزمن والانعكاس فى المعامل.

المحتويات

الباب الأول

المفاهيم الاحصائية ٧

الباب الثانى

أساليب عرض البيانات ٢٥

الباب الثالث

المقاييس الاحصائية للبيانات ١٥٩

الباب الرابع

الارتباط الخطى ٢٦٥

الباب الخامس

الانحدار الخطى ٣٠٢

الباب السادس

صناعة السياحة فى مصر ٣٣٩





مكتبة بستان المعرفة

لطباعة ونشر وتوزيع الكتب

كفر الدوار – الحدائق – بجوار نقابة التطبيقيين

☎ ٠٤٥/٢٢٢٤٢٢٨ الإسكندرية: ٠١٢٢٣٥٣٤٨١٤